

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 18:11:59 ; Search time 5268 Seconds
(without alignments)
11082.587 Million cell updates/sec

Title: US-10-650-369-21

Perfect score: 1347
Sequence: 1 atgaaaaataacagggat.....ttgcaaaatcgctaataa 1347

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.cm.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.roi.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.roi.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.tod.*

36: em.htg.nam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htg.hum.*

40: em.htg.mus.*

41: em.htg.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1347	100.0	1347	6	AR437352	AX37352 Sequence
2	1347	100.0	1347	6	AX367176	AX367176 Sequence
3	817.4	60.7	1011	1	AP375662	AP375662 Streptococ
4	815.8	60.6	1011	6	AR437347	AR437347 Sequence
5	815.8	60.6	1011	6	AX355986	AX355986 Sequence
6	815.8	60.6	1011	6	AX367166	AX367166 Sequence
7	782.2	58.1	1158	1	AB088216	AB088216 Streptococ
8	782.2	58.1	1158	1	AB088217	AB088217 Streptococ
9	782.2	58.1	1158	1	AB088214	AB088214 Streptococ
10	780.6	58.0	1104	1	AB088215	AB088215 Streptococ
11	780.6	58.0	1125	6	AR135510	AR135510 Sequence
12	780.6	58.0	1156	1	STRPLR	M95569 Streptococ
13	780.6	58.0	2326	1	AF520813	AF520813 Streptococ
14	780.6	58.0	10053	1	AE006494	AE006494 Streptococ
15	780.6	58.0	10029	1	AE014140	AE014140 Streptococ
16	780.6	58.0	311600	1	AP005141	AP005141 Streptococ
17	779	57.8	12001	1	AE009973	AE009973 Streptococ
18	757.6	56.2	1260	1	SEGAPCGN	X57786 S.equismil
19	757.6	56.2	2334	1	SEGAPLPP	Y12602 Streptococ
20	702.2	52.1	1011	6	AR437348	AR437348 Sequence
21	702.2	52.1	1011	6	AX355988	AX355988 Sequence
22	702.2	52.1	1011	6	AX367168	AX367168 Sequence
23	700.6	52.0	1011	6	AX607902	AX607902 Sequence
24	700.6	52.0	12960	6	AX602161	AX602161 Sequence
25	700.6	52.0	122050	1	SAG766853	AL766853 Streptococ
26	699	51.9	1011	1	AF338416	AF338416 Streptococ
27	699	51.9	21954	1	AE014272	AE014272 Streptococ
28	694.2	51.5	1092	1	AB110908	AB110908 Streptococ
29	686.2	50.9	1012	1	AY167026	AY167026 Streptococ
30	683	50.7	7034	2	SPNBU1925	AL449947 Streptococ
31	678.2	50.3	1008	1	SPN305822	AU505822 Streptococ
32	678.2	50.3	1011	1	AF421899	AF421899 Streptococ
33	678.2	50.3	1011	1	AF421902	AF421902 Streptococ
34	678.2	50.3	1011	6	AR437351	AR437351 Sequence
35	678.2	50.3	1011	6	AX355994	AX355994 Sequence
36	678.2	50.3	1011	6	AX367174	AX367174 Sequence
37	678.2	50.3	7571	6	AR218770	AR218770 Sequence
38	678.2	50.3	7571	6	BD003682	BD003682 Polynucle
39	678.2	50.3	10145	1	AE007490	AE007490 Streptococ
40	678.2	50.3	12139	1	AE008547	AE008547 Streptococ
41	678.2	50.3	349980	6	AX571765	AX571765 Sequence
42	678.2	50.3	349980	6	AX571766	AX571766 Sequence
43	676.6	50.2	1011	1	AF421900	AF421900 Streptococ
44	676.6	50.2	1011	6	AR437349	AR437349 Sequence
45	676.6	50.2	1011	6	AX355990	AX355990 Sequence

ALIGNMENTS

RESULT 1	AR437352	Sequence 21 from patent US 6660270.	1347 bp	DNA	linear	PAT 18-DEC-2003
AR437352	AR437352	Sequence 21 from patent US 6660270.	1347 bp	DNA	linear	PAT 18-DEC-2003
LOCUS	AR437352	Sequence 21 from patent US 6660270.	1347 bp	DNA	linear	PAT 18-DEC-2003
DEFINITION	AR437352	Sequence 21 from patent US 6660270.	1347 bp	DNA	linear	PAT 18-DEC-2003
ACCESSION	AR437352	Sequence 21 from patent US 6660270.	1347 bp	DNA	linear	PAT 18-DEC-2003
VERSION	AR437352.1	GI:40202283	1347 bp	DNA	linear	PAT 18-DEC-2003
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 1347)					
AUTHORS	Potter, A.A., Perez-Casal, J. and Fontaine, M.					
TITLE	Immunization of dairy cattle with chimeric GapC protein against					
JOURNAL	Streptococcus infection					
	Patent: US 6660270-A 21 09-DEC-2003;					

[illegible]

QY 1 ATGAAAAAATAACAGGATATATTTATTTGCTCTTTCAGTCAATATTTCTGTCTGATGC 60
Db 1 ATGAAAAAATAACAGGATATATTTATTTGCTCTTTCAGTCAATATTTCTGTCTGATGC 60
QY 61 CAGGCAAACTACGGATCCGGTATGTTAAAGTTGGTATTAACGGTTTCGGTCTGATC 120
Db 61 CAGGCAAACTACGGATCCGGTATGTTAAAGTTGGTATTAACGGTTTCGGTCTGATC 120
QY 121 GAGCGTTGCTCATTCGGTCTGATTAACAAATGTTGAAGGTTGTTGAAGTAACTCGTATCAAC 180
Db 121 GAGCGTTGCTCATTCGGTCTGATTAACAAATGTTGAAGGTTGTTGAAGTAACTCGTATCAAC 180
QY 181 GACCTTACAGATCCAAACATGCTTGCACACTGTTGCAATACGATCACTCAAGAGCT 240
Db 181 GACCTTACAGATCCAAACATGCTTGCACACTGTTGCAATACGATCACTCAAGAGCT 240
QY 241 TTGACGGAACCTTGAAGTTAAAGAGGTTGATTTGAAGTAAACGGAATTCATCAAA 300
Db 241 TTGACGGAACCTTGAAGTTAAAGAGGTTGATTTGAAGTAAACGGAATTCATCAAA 300
QY 301 GTTCTGCTGACGATCCAGAAACATCGACTGGGCACTGACGGTGTGAAATCGTT 360
Db 301 GTTCTGCTGACGATCCAGAAACATCGACTGGGCACTGACGGTGTGAAATCGTT 360
QY 361 CTGGAAGCACTCGAGGTTACTGTAGAAGTTAAAGATGTTGATTTGACGTTAAACGGA 420
Db 361 CTGGAAGCACTCGAGGTTACTGTAGAAGTTAAAGATGTTGATTTGACGTTAAACGGA 420
QY 421 TTCAATTAAGTTTCTGCTGAAAGATCCAGAAACATGCTGGGCACTGACGGTGT 480
Db 421 TTCAATTAAGTTTCTGCTGAAAGATCCAGAAACATGCTGGGCACTGACGGTGT 480
QY 481 GAAATCGTTCTTGAATCGATGCTGTTGAAGTTAAAGAGGTTGATTTGAAGTAAAC 540
Db 481 GAAATCGTTCTTGAATCGATGCTGTTGAAGTTAAAGAGGTTGATTTGAAGTAAAC 540
QY 541 GGTCAATTTGTTAAAGTTTCTGCTGAAAGATCCAGAAACATGCTGGGCACTGAT 600
Db 541 GGTCAATTTGTTAAAGTTTCTGCTGAAAGATCCAGAAACATGCTGGGCACTGAT 600
QY 601 GCGGTAGAATCGTTCTGCAAGCACTAGTTCTTGTGTTAAAGAGGTTGATTTGAAG 660
Db 601 GCGGTAGAATCGTTCTGCAAGCACTAGTTCTTGTGTTAAAGAGGTTGATTTGAAG 660
QY 661 CACTTACATGCTAAACGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 CACTTACATGCTAAACGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 AAAACAGTTGTTTCAACACTAACCGACATCTTGAAGGTTGATTTGAAGTAAAC 780
Db 721 AAAACAGTTGTTTCAACACTAACCGACATCTTGAAGGTTGATTTGAAGTAAAC 780
QY 781 GGTGCTTCATGCTACAACTGTTTGTAGCTCTATGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 781 GGTGCTTCATGCTACAACTGTTTGTAGCTCTATGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 841 GGTATCCAAAAGGTTCTTATGCTACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 841 GGTATCCAAAAGGTTCTTATGCTACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 901 GACGACCAACCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 901 GACGACCAACCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 961 CTTAACTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 961 CTTAACTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 CTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 CTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 ACTCTTGATAAAAGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140

Db 1081 ACTCTTGATAAAAGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1141 GACAGTTTCGGTTACACTGAAGATCCAAATGTTTCTTCCAGATATCGTAGGCGTGTCTAC 1200
Db 1141 GACAGTTTCGGTTACACTGAAGATCCAAATGTTTCTTCCAGATATCGTAGGCGTGTCTAC 1200
QY 1201 GGTTCATTTGTTGACGCAACTCAAACTAAAGTTATGGAAGTTGACGATCAAAATGTT 1260
Db 1201 GGTTCATTTGTTGACGCAACTCAAACTAAAGTTATGGAAGTTGACGATCAAAATGTT 1260
QY 1261 AAAGTTGATCATGTTGATGATCAATGAATGTTTCTTACACTGCTGCTGCTGCTGCTGCT 1320
Db 1261 AAAGTTGATCATGTTGATGATCAATGAATGTTTCTTACACTGCTGCTGCTGCTGCTGCT 1320
QY 1321 GAGTATTTTGCAAAATCGCTAAATAA 1347
Db 1321 GAGTATTTTGCAAAATCGCTAAATAA 1347
RESULT 3
AF375662 1011 bp DNA linear BCT 03-MAY-2003
LOCUS Streptococcus dysgalactiae glycerolaldehyde 3-P dehydrogenase (gapC)
DEFINITION gene, complete cds.
ACCESSION AF375662
VERSION AF375662.1 GI:30348860
SOURCE Streptococcus dysgalactiae
ORGANISM Streptococcus dysgalactiae
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 1011)
Perez-Casal,J.F., Bolton,A., Song,X.-M., Willson,P. and Potter,A.A.
AUTHORS Use of the surface proteins GapC and Mig of Streptococcus
TITLE dysgalactiae as protective antigens against mastitis in
JOURNAL non-lactating cows
REFERENCE 2 (bases 1 to 1011)
Perez-Casal,J.F., Bolton,A., Song,X.-M., Willson,P. and Potter,A.A.
AUTHORS Direct Submission
TITLE Submitted (02-MAY-2001) VIDO, University of Saskatchewan, 120
JOURNAL Veterinary Rd, Saskatoon, SK S7N 5E3, Canada
FEATURES
source
1..1011
/organism="Streptococcus dysgalactiae"
/mol_type="genomic DNA"
/db_xref="taxon:1334"
gene
1..1011
/gene="gapC"
CDS
1..1011
/genes="gapC"
/notes="GapC"
/codon_start=1
/transl_table=11
/product="glycerolaldehyde 3-P dehydrogenase"
/protein_id="AAP31408.1"
/db_xref="GI:30348861"
LKYDTTQGRFDGTVVEKGGFVNGNFIKVSABRDPENDMDATDGVIEI VLEATGFFAK
KEAEKHLHANGAKKVIITAPGNDVKVTVPVTHNHLIDGTEVIVSGACTTNC LAPM
AKALHPAFGIQGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAIIVPNSITGA AKAI
GLVIBELNGKLDGAQRVPVPTGTVTELVDLKNVSDIENAAKKAANDSFGYTED
PIVSSDIVGSYSLFDATQTKYMEVDGSQLVKVWSYDNEWSYTAQLVRLTEYFAKI
AK"
ORIGIN
Query Match 50.7%; Score 817.4; DB 1; Length 1011;
Best Local Similarity 97.5%; Pred. NO. 1.1e-154; Indels 0; Gaps 0;
Matches 830; Conservative 0; Mismatches 21;
QY 497 TCGATGCTACTGTTGAAGTTAAAGAGGTTGATTTGAAGTAACTGCTGCTGCTGCTGCT 556

Db 161 TTGACGGAACTGTTGAAGTTAAAGAGGTGGATTTTGAAGTAAACCGAAACCTTCATCAAG 220
 QY 557 TTTCTGCTGAACGCAACAGCAACATGACTGGCTACTGATGGCGTAGAAATCGTTC 616
 Db 221 TTTCTGCTGAACGCTGATCAGAAACATCGACTGGGCAACTGACGGTGTGAATCGTTC 280
 QY 617 TTGAAGCACTAGTTCTTTGCTAAAGAGAGCTGCTGAAGAACACATTCATGCTAACG 676
 Db 281 TGAAGCACTAGTTCTTTGCTAAAGAGAGCTGCTGAAGAACACATTCATGCTAACG 340
 QY 677 GTGCTAAAGAAAGTTGTTATCAGCTCCTGTGGGAAACGAGTTTAAACAGTGTGTTTCA 736
 Db 341 GTGCTAAAGAAAGTTGTTATCAGCTCCTGTGGGAAACGAGTTTAAACAGTGTGTTTCA 400
 QY 737 ACACCTAACCCAGCACTTCTGACGCTACTGAAACAGTTATCTCAGGTCCTTCATGCTACTA 796
 Db 401 ACACCTAACCCAGCACTTCTGACGCTACTGAAACAGTTATCTCAGGTCCTTCATGCTACTA 460
 QY 797 CAAACTGTTTGTAGCTCCTATGGCTAAAGCTCTTACAGATGCAATTTGGTATCCAAAAGGTC 856
 Db 461 CAAACTGTTTGTAGCTCCTATGGCTAAAGCTCTTACAGATGCAATTTGGTATCCAAAAGGTC 520
 QY 857 TTATGACTAACCTCAGCGTTATCTGTGACCAAAATGATCCTTGACGGACCAACCGTG 916
 Db 521 TTATGACTAACCTCAGCGTTATCTGTGACCAAAATGATCCTTGACGGACCAACCGTG 580
 QY 917 GTGGTGACCTTCGTCGCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
 Db 581 GTGGTGACCTTCGTCGCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
 QY 977 GTGCTAAAGAAAGTTGTTATCAGCTCCTGTGGGAAACGAGTTTAAACAGTGTGTTTCA 1036
 Db 641 GTGCTAAAGAAAGTTGTTATCAGCTCCTGTGGGAAACGAGTTTAAACAGTGTGTTTCA 700
 QY 1037 AACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1096
 Db 701 AACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 760
 QY 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAACGACAGTTTCGGTTTACA 1156
 Db 761 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAACGACAGTTTCGGTTTACA 820
 QY 1157 CTGAAGTCCAAATGTTTCTTCAATATCGTAGCGGTGCTACAGGTTCAATGTTTGACG 1216
 Db 821 CTGAAGTCCAAATGTTTCTTCAATATCGTAGCGGTGCTACAGGTTCAATGTTTGACG 880
 QY 1217 CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTAAAGTTTGTATCATGGT 1276
 Db 881 CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTAAAGTTTGTATCATGGT 940
 QY 1277 ATGCAATGAATGCTTTACACTGCTCAACTGTTGCTACACTTGAGTATTTTGCAAAA 1336
 Db 941 ATGCAATGAATGCTTTACACTGCTCAACTGTTGCTACACTTGAGTATTTTGCAAAA 1000
 QY 1337 TCGCTAAATAA 1347
 Db 1001 TCGCTAAATAA 1011

RESULT 4
 AR437347
 LOCUS 1011 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 11 from patent US 6660270.
 ACCESSION AR437347
 VERSION AR437347.1 GI:40202278
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1011)
 AUTHORS Porter, A., Perez-Casal, J. and Fontaine, M.
 TITLE Immunization of dairy cattle with chimeric GapC protein against Streptococcus infection

JOURNAL Patent: US 6660270-A 11 09-DEC-2003;

FEATURES
 Location/Qualifiers
 1..1011

source
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 60.6%; Score 815.8; DB 6; Length 1011;
 Best Local Similarity 97.4%; Pred. No. 2,3e-154;
 Matches 829; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 497 TCATCGTACTGTTGAAGTTAAAGAGGTGGATTTTGAAGTAAACCGAAACCTTCATCAAG 556
 Db 161 TTGACGGAACTGTTGAAGTTAAAGAGGTGGATTTTGAAGTAAACCGAAACCTTCATCAAG 220
 QY 557 TTTCTGCTGAACGCGAACCCAGCAACATTCAGTCTGCGCTACTGATGGCGTAGAAATCGTTC 616
 Db 221 TTTCTGCTGAACGCTGATCCAGAAACATCGACTGGGCAACTGACGGTGTGAATCGTTC 280
 QY 617 TTGAAGCACTAGTTCTTTGCTAAAGAGAGCTGCTGAAGAACACATTCATGCTAACG 676
 Db 281 TGAAGCACTAGTTCTTTGCTAAAGAGAGCTGCTGAAGAACACATTCATGCTAACG 340
 QY 677 GTGCTAAAGAAAGTTGTTATCAGCTCCTGTGGGAAACGAGTTTAAACAGTGTGTTTCA 736
 Db 341 GTGCTAAAGAAAGTTGTTATCAGCTCCTGTGGGAAACGAGTTTAAACAGTGTGTTTCA 400
 QY 737 ACACCTAACCCAGCACTTCTGACGCTACTGAAACAGTTATCTCAGGTCCTTCATGCTACTA 796
 Db 401 ACACCTAACCCAGCACTTCTGACGCTACTGAAACAGTTATCTCAGGTCCTTCATGCTACTA 460
 QY 797 CAAACTGTTTGTAGCTCCTATGGCTAAAGCTCTTACAGATGCAATTTGGTATCCAAAAGGTC 856
 Db 461 CAAACTGTTTGTAGCTCCTATGGCTAAAGCTCTTACAGATGCAATTTGGTATCCAAAAGGTC 520
 QY 857 TTATGACTAACCTCAGCGTTATCTGTGACCAAAATGATCCTTGACGGACCAACCGTG 916
 Db 521 TTATGACTAACCTCAGCGTTATCTGTGACCAAAATGATCCTTGACGGACCAACCGTG 580
 QY 917 GTGGTGACCTTCGTCGCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
 Db 581 GTGGTGACCTTCGTCGCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
 QY 977 CTGCTAAAGAAAGTTGTTATCAGCTCCTGTGGGAAACGAGTTTAAACAGTGTGTTTCA 1036
 Db 641 CTGCTAAAGAAAGTTGTTATCAGCTCCTGTGGGAAACGAGTTTAAACAGTGTGTTTCA 700
 QY 1037 AACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1096
 Db 701 AACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 760
 QY 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAACGACAGTTTCGGTTTACA 1156
 Db 761 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAACGACAGTTTCGGTTTACA 820
 QY 1157 CTGAAGTCCAAATGTTTCTTCAATATCGTAGCGGTGCTACAGGTTCAATGTTTGACG 1216
 Db 821 CTGAAGTCCAAATGTTTCTTCAATATCGTAGCGGTGCTACAGGTTCAATGTTTGACG 880
 QY 1217 CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTAAAGTTTGTATCATGGT 1276
 Db 881 CAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTAAAGTTTGTATCATGGT 940
 QY 1277 ATGCAATGAATGCTTTACACTGCTCAACTGTTGCTACACTTGAGTATTTTGCAAAA 1336
 Db 941 ATGCAATGAATGCTTTACACTGCTCAACTGTTGCTACACTTGAGTATTTTGCAAAA 1000
 QY 1337 TCGCTAAATAA 1347
 Db 1001 TCGCTAAATAA 1011

RESULT 5


```

AX355986
LOCUS AX355986 1011 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 3 from Patent WO0196381.
ACCESSION AX355986
VERSION AX355986.1 GI:18620580
KEYWORDS Streptococcus dysgalactiae
SOURCE Streptococcus dysgalactiae
ORGANISM Streptococcus dysgalactiae
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
1
REFERENCE
AUTHORS Bolton,A.J., Perez-Casal,J., Fontaine,M. and Potter,A.A.
TITLE Immunization of dairy cattle with gapc protein against
Streptococcus infection
JOURNAL Patent: WO 0196381-A 3 20-DEC-2001;
The University of Saskatchewan (CA)
FEATURES
source
1..1011
/organism="Streptococcus dysgalactiae"
/mol_type="unassigned DNA"
/db_xref="taxon:1334"
1..1011
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23006.1"
/db_xref="GI:18620581"
/db_xref="RENTREMBL:CAD23006"
/translation="MVVKVINGFGFRIGRLAFRRIONVEGVETRLNDLPNMLAHL
LKDYDTGFRDVEVKEGFEVNGFNFKVSAERDENIDWATDGVLEIVLEATGFFAK
KEAALHANGAKKVIITAPGNDVKVTFVFNTHDILDGTETVSGASCTTNCALPM
AKALHDFGIGKGLMTTHAYTGDDMLQDHRGDLRRARAGAAIIVPNSGAAKAI
GLVIPLENGKLDGAAQRPVPTGVSVELVTLKDVSVDEINAAKAAASNDSEGYTED
PIVSSDIVGSVGSFLDQTQKMEVDGSQLVKVSWYDNENSYTAQLVRLLEYFAKI
AK"
ORIGIN
Query Match 60.6%; Score 815.8; DB 6; Length 1011;
Best Local Similarity 97.4%; Pred. No. 2.3e-154;
Matches 829; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 497 TCGATGGTACTGTGAAGTTAAGAAGTGGATCGAAGTTAACGGTCAATTTGTTAAAG 556
Db 161 TTGACGGAAGTGTGAAGTTAAGAAGTGGATCGAAGTTAACGGTCAATTTGTTAAAG 220
QY 557 TTCTCTGCTGAACCGCAACAGCAAAACATTGACTGGGCTACTGATGCGTAGAAATCGTTC 616
Db 221 TTCTCTGCTGAACCGCAACAGCAAAACATTGACTGGGCTACTGATGCGTAGAAATCGTTC 280
QY 617 TTGAAGCAACTAGTTTCTTTGCTTAAAGAGAGAGCTGCTGAAAAACACATTACATGCTAACG 676
Db 281 TGAAGCAACTAGTTTCTTTGCTTAAAGAGAGAGCTGCTGAAAAACACATTACATGCTAACG 340
QY 677 GTGCTAAAAAGTTGTTATACAGCTCCTGGTGGAAACGACGTTAAACAGTTGTTTCA 736
Db 341 GTGCTAAAAAGTTGTTATACAGCTCCTGGTGGAAACGACGTTAAACAGTTGTTTCA 400
QY 737 ACACTAACCAACGACATTCTTTCAGGCTACTGAAACAGTTATCTCAGTGCTTCATGCTACTA 796
Db 401 ACACTAACCAACGACATTCTTTCAGGCTACTGAAACAGTTATCTCAGTGCTTCATGCTACTA 460
QY 797 CAACTGTTTACCTCTATGCTTAAGCTCTTCAGATGCTATTCGATTCACAAAAGGTC 856
Db 461 CAACTGTTTACCTCTATGCTTAAGCTCTTCAGATGCTATTCGATTCACAAAAGGTC 520
QY 857 TTATGACTACAATCCACCGCTTATCTGTTGACCAATGATCTTCGACGACACACACCGGTG 916
Db 521 TTATGACTACAATCCACCGCTTATCTGTTGACCAATGATCTTCGACGACACACACCGGTG 580
QY 917 GTGGTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 581 GTGGTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640

LOCUS AX367166 1011 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 11 from Patent WO0196379.
ACCESSION AX367166
VERSION AX367166.1 GI:18855364
KEYWORDS Streptococcus dysgalactiae
SOURCE Streptococcus dysgalactiae
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
1
REFERENCE
AUTHORS Potter,A.A., Perez-Casal,J. and Fontaine,M.
TITLE Immunization of dairy cattle with chimeric gapc protein against
Streptococcus infection
JOURNAL Patent: WO 0196379-A 11 20-DEC-2001;
The University of Saskatchewan (CA)
FEATURES
source
1..1011
/organism="Streptococcus dysgalactiae"
/mol_type="unassigned DNA"
/db_xref="taxon:1334"
1..1011
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23838.1"
/db_xref="GI:18855365"
/db_xref="RENTREMBL:CAD23838"
/translation="MVVKVINGFGFRIGRLAFRRIONVEGVETRLNDLPNMLAHL
LKDYDTGFRDVEVKEGFEVNGFNFKVSAERDENIDWATDGVLEIVLEATGFFAK
KEAALHANGAKKVIITAPGNDVKVTFVFNTHDILDGTETVSGASCTTNCALPM
AKALHDFGIGKGLMTTHAYTGDDMLQDHRGDLRRARAGAAIIVPNSGAAKAI
GLVIPLENGKLDGAAQRPVPTGVSVELVTLKDVSVDEINAAKAAASNDSEGYTED
PIVSSDIVGSVGSFLDQTQKMEVDGSQLVKVSWYDNENSYTAQLVRLLEYFAKI
AK"
ORIGIN
Query Match 60.6%; Score 815.8; DB 6; Length 1011;
Best Local Similarity 97.4%; Pred. No. 2.3e-154;
Matches 829; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 977 CTGCTAAAGCTATCGGTTCTTTATCCAGAAATGAATGGTAAACTTGTATGTTGCTGCAC 1036
Db 641 CTGCTAAAGCTATCGGTTCTTTATCCAGAAATGAATGGTAAACTTGTATGTTGCTGCAC 700
QY 1037 AACGTGTTCTGTTCCCACTCGATCAGTAACTGAGTTGGTTGTAATCTTGTATGATAAAG 1096
Db 701 AACGTGTTCTGTTCCCACTCGATCAGTAACTGAGTTGGTTGTAATCTTGTATGATAAAG 760
QY 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTTTCGTTTACA 1156
Db 761 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTTTCGTTTACA 820
QY 1157 CTGAAGATCCAATTTGTTCTTTCAGATATCGTAGCGGTGTCATACGGTTCAATGTTGACG 1216
Db 821 CTGAAGATCCAATTTGTTCTTTCAGATATCGTAGCGGTGTCATACGGTTCAATGTTGACG 880
QY 1217 CAACCTCAAACTAAAAGTTATGGAAGTTGACCGATCACAATCGTTTAAAGTTGATCATCGGT 1276
Db 881 CAACCTCAAACTAAAAGTTATGGAAGTTGACCGATCACAATCGTTTAAAGTTGATCATCGGT 940
QY 1277 ATGACATGAAATGCTTTACACTGCTCAACTGTTGCTGACACTTGAATTTTCAAAAA 1336
Db 941 ATGACATGAAATGCTTTACACTGCTCAACTGTTGCTGACACTTGAATTTTCAAAAA 1000
QY 1337 TCGCTAAATAA 1347
Db 1001 TCGCTAAATAA 1011

RESULT 6
LOCUS AX367166 1011 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 11 from Patent WO0196379.
ACCESSION AX367166
VERSION AX367166.1 GI:18855364
KEYWORDS Streptococcus dysgalactiae
SOURCE Streptococcus dysgalactiae
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
1
REFERENCE
AUTHORS Potter,A.A., Perez-Casal,J. and Fontaine,M.
TITLE Immunization of dairy cattle with chimeric gapc protein against
Streptococcus infection
JOURNAL Patent: WO 0196379-A 11 20-DEC-2001;
The University of Saskatchewan (CA)
FEATURES
source
1..1011
/organism="Streptococcus dysgalactiae"
/mol_type="unassigned DNA"
/db_xref="taxon:1334"
1..1011
/notes="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23838.1"
/db_xref="GI:18855365"
/db_xref="RENTREMBL:CAD23838"
/translation="MVVKVINGFGFRIGRLAFRRIONVEGVETRLNDLPNMLAHL
LKDYDTGFRDVEVKEGFEVNGFNFKVSAERDENIDWATDGVLEIVLEATGFFAK
KEAALHANGAKKVIITAPGNDVKVTFVFNTHDILDGTETVSGASCTTNCALPM
AKALHDFGIGKGLMTTHAYTGDDMLQDHRGDLRRARAGAAIIVPNSGAAKAI
GLVIPLENGKLDGAAQRPVPTGVSVELVTLKDVSVDEINAAKAAASNDSEGYTED
PIVSSDIVGSVGSFLDQTQKMEVDGSQLVKVSWYDNENSYTAQLVRLLEYFAKI
AK"
ORIGIN
Query Match 60.6%; Score 815.8; DB 6; Length 1011;
Best Local Similarity 97.4%; Pred. No. 2.3e-154;
Matches 829; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 497 TCGATGGTACTGTGAAGTTAAGAAGTGGATCGAAGTTAACGGTCAATTTGTTAAAG 556
Db 161 TTGACGGAAGTGTGAAGTTAAGAAGTGGATCGAAGTTAACGGTCAATTTGTTAAAG 220
QY 557 TTCTCTGCTGAACCGCAACAGCAAAACATTGACTGGGCTACTGATGCGTAGAAATCGTTC 616
Db 221 TTCTCTGCTGAACCGCAACAGCAAAACATTGACTGGGCTACTGATGCGTAGAAATCGTTC 280
QY 617 TTGAAGCAACTAGTTTCTTTGCTTAAAGAGAGAGCTGCTGAAAAACACATTACATGCTAACG 676
Db 281 TGAAGCAACTAGTTTCTTTGCTTAAAGAGAGAGCTGCTGAAAAACACATTACATGCTAACG 340
QY 677 GTGCTAAAAAGTTGTTATACAGCTCCTGGTGGAAACGACGTTAAACAGTTGTTTCA 736
Db 341 GTGCTAAAAAGTTGTTATACAGCTCCTGGTGGAAACGACGTTAAACAGTTGTTTCA 400
QY 737 ACACTAACCAACGACATTCTTTCAGGCTACTGAAACAGTTATCTCAGTGCTTCATGCTACTA 796
Db 401 ACACTAACCAACGACATTCTTTCAGGCTACTGAAACAGTTATCTCAGTGCTTCATGCTACTA 460
QY 797 CAACTGTTTACCTCTATGCTTAAGCTCTTCAGATGCTATTCGATTCACAAAAGGTC 856
Db 461 CAACTGTTTACCTCTATGCTTAAGCTCTTCAGATGCTATTCGATTCACAAAAGGTC 520
QY 857 TTATGACTACAATCCACCGCTTATCTGTTGACCAATGATCTTCGACGACACACACCGGTG 916
Db 521 TTATGACTACAATCCACCGCTTATCTGTTGACCAATGATCTTCGACGACACACACCGGTG 580
QY 917 GTGGTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 581 GTGGTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640

```


Db	161	TTGAGCGNACTGTTGAAGCTTAAAGAGAGTGGATTTGAAGTAAACGGAAACTTCATCAAG	220
Qy	557	TTTCTGTGTAACGGGAACCAACAAACATTTGACTGGCTTACTGATGGGTAGAAATCGTTC	615
Db	221	TTTCTGTGTAACGCTGATCCAGAAACATCGACTGGCAACTGACGGTGTTCGAATCGTTC	280
Qy	617	TTGAAGCAACTAGTTCCTTTCTCTAAAAAAGAGCTGCTGAAAAACACTTACATGCTCAACG	675
Db	281	TGGAAGCAACTGGTTCCTTTGCTTAAAAAGAAAGCTGCTGAAAAACATTTACATGCTCAACG	340
Qy	677	GTGCTAAAAAGTTGTGTATATCAACAGCTCTCTGTGTGAAACGACGTTAAACAGTGTTCATCA	736
Db	341	GTGCTAAAAAGTTGTGTATATCAACAGCTCTCTGTGTGAAACGACGTTAAACAGTGTTCATCA	400
Qy	737	ACACTAACACGACATTCCTTGACGGTACTGAAAAAGATTATCTCAGGTGCTTCATGTACTA	799
Db	401	ACACTAACACGACATTCCTTGACGGTACTGAAAAAGATTATCTCAGGTGCTTCATGTACTA	460
Qy	797	CAAACTGTTAGCTCCCTATGGCTAAAGCTCTTCCAGATGCATTTGGTATCCAAAAAGGTC	856
Db	461	CAAACTGTTAGCTCCCTATGGCTAAAGCTCTTCCAGATGCATTTGGTATCCAAAAAGGTC	520
Qy	857	TTATGACTCAAAATCCAGCTTATACTGGTGACCAAATGATCCTTTGACGGACCAACCGGTG	916
Db	521	TTATGACTCAAAATCCAGCTTATACTGGTGACCAAATGATCCTTTGACGGACCAACCGGTG	580
Qy	917	GTGGTGACCTTCGTCGGTCTCGTGGTGTGCTGCAAAACATTTGTTCCCTAACTCAACTGGTG	976
Db	581	GTGGTGACCTTCGTCGGTCTCGTGGTGTGCTGCAAAACATTTGTTCCCTAACTCAACTGGTG	640
Qy	977	CTGCTAAAGCTATCCGGTCTTGTTATCCAGAAATTTGAATGGTTAAACTTTGATGGTCTGCAC	1036
Db	641	CTGCTAAAGCTATCCGGTCTTGTTATCCAGAAATTTGAATGGTTAAACTTTGATGGTCTGCAC	700
Qy	1037	AAGGTGTTCTGTCCAACTGGATCAGTAACTGATGGTTGGTTGTAATCTTTGATAAAAACG	1096
Db	701	AAGGTGTTCTGTCCAACTGGATCAGTAACTGATGGTTGGTTGTAATCTTTGATAAAAACG	760
Qy	1097	TTTCTGTTGACGAATCAACGCTGCTATGAAGCTGCTTCAACGACAGATTTCCGGTTACA	1156
Db	761	TTTCTGTTGACGAATCAACGCTGCTATGAAGCTGCTTCAACGACAGATTTCCGGTTACA	820
Qy	1157	CTGAAGATCAAAATGTTTCTTCAGATATCGTAGCGGTGCATACGGTTCATTTGTTGAACG	1216
Db	821	CTGAAGATCAAAATGTTTCTTCAGATATCGTAGCGGTGCATACGGTTCATTTGTTGAACG	880
Qy	1217	CAAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTTAAAGTTGTATCATGGT	1276
Db	881	CAAACTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTTAAAGTTGTATCATGGT	940
Qy	1277	ATGACATGAAGATGCTTTTACACTGCTCAAATTTGTTTGGTACACTTGAGTATTTTGCRAAAA	1336
Db	941	ATGACATGAAGATGCTTTTACACTGCTCAAATTTGTTTGGTACACTTGAGTATTTTGCRAAAA	1000
Qy	1337	TCGCTAAATAA 1347	
Db	1001	TCGCTAAATAA 1011	

RESULT	7
AB088216	
LOCUS	1158 bp DNA linear BCT 27-JUL-2002
DEFINITION	Streptococcus pyogenes napr gene for nephritis-associated plasmin receptor, complete cds, strain:M type 12.
ACCESSION	AB088216
VERSION	AB088216.1 GI:21886629
KEYWORDS	
SOURCE	Streptococcus pyogenes
ORGANISM	Streptococcus pyogenes
	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
	Streptococcus.
REFERENCE	1
AUTHORS	Yoshizawa, N., Yamakami, K. and Fujino, M.

```

TITLE      Further studies of nephritis-associated plamin receptor (NAPlr) in
REFERENCE  acute poststreptococcal glomerulonephritis (APSGN)
JOURNAL    Unpublished
AUTHORS    2 (bases 1 to 1158)
           Fujino,M., Yamakami,K. and Yoshizawa,N.
TITLE      Direct Submission
JOURNAL    Submitted (15-JUL-2002) Nobuyuki Yoshizawa, national defense
           medical college, department of public health; 3-2 namiki,
           tokorozawa, saitama 359-8513, Japan (E-mail:yosizawa@cc.ndmc.ac.jp)
           Tel:81-42-995-1575, Fax:81-42-996-5196)
FEATURES   Location/Qualifiers
           1..1158
             /organism="Streptococcus pyogenes"
             /mol_type="genomic DNA"
             /strain="M type 12"
             /db_xref="taxon:1314"
             13..18
             35..40
             100..106
             117..1127
             /genes="naplr"
             117..1127
             /genes="naplr"
             /EC_number="1.2.1.12"
             /function="glyceraldehyde-3-phosphate dehydrogenase;
             plasmn binding activity"
             /notes="glyceraldehyde-3-phosphate dehydrogenase (GAPDH);
             plasmn binding protein; nephritogenic anti:gen"
             /codon_start=1
             /evidence=experimental
             /transl_table=11
             /product="nephritis-associated plasmn receptor"
             /protein_id="BAC05496.1"
             /db_xref="GI:21886630"
             /translation="MVVKYINGFGIRGLAPRIQIEGVETRLDLPNNMLHL
             LKYTEQGFEDTVEKEGFVNGFIKVSBERDSENIDMA2DGEIVLEATGFPFAK
             KEAEKHLHANGAKVVTIAPGNDVKTUVFVNNHDLIGTETVIGASCTTNCLEAPM
             AKALHDHAFQIQLMHTIYATGDQIMLDGPHRGDRLRRAGAAIVPNSTGAAKAI
             GLVIPELNGKLGAATFVGTSLVTLTDKNVSDVEIPNAAMKAASNDSGFYED
             PTVSSDITGVSGYSLFDATQTKMVEDGSQLVKVSVWYDNNESYTAQLVRLTEYFAK
             AK"

```

ORIGIN

Query Match	58.1%	Score 782.2	DB 1	Length 1158
Best Local Similarity	94.9%	Pred. No. 1.4e-147		
Matches 808	Conservative 0	Mismatches 43	Indels 0	Gaps 0
QY	497	TCGATCGTACTGTTGAAGTTAAAGACAGGTGGATTCCGAAGTTAACGCTCAATTGTTAAAG	556	
Db	277	TTGATGGAACAGTTGAAGTTAAAGAGGTGGATTTCGAAGTTAAACGGAACCTTCATTAAAG	336	
QY	557	TTTCTGCTGAACCGCAACACAGCAAAACATTGACGTGGGCTACTGATGGCGTAGAAATCGTTC	616	
Db	337	TTTCTGCTGACGTGATCAGAAAAACATCGACTGGCGCACTGATGGGGTTGAATCGTTC	396	
QY	617	TTGAAGCAACTAGTTCTTTTGTAAAAAAGAAGCTCTCTGAAAAACACTTACATGCTAAAG	676	
Db	397	TTGAAGCAACTGTTCTTTTGTCTAAAAAAGAAGCAGCTGAAAAACACTTACATGCTAAAG	456	
QY	677	GTGCTAAAAAAGTTGTATACAGGCTCCTGTGTGGAAACGACGTTAAACAGTGTGTTTCA	736	
Db	457	GTGCTAAAAAAGTTGTATACAGGCTCCTGTGTGGAAACGATGTAAACAGTGTGTTTCA	516	
QY	737	ACACTAACACGACATCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTTCATGTACTA	796	
Db	517	ACACTAACACGACATCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTTCATGTACTA	576	
QY	797	CAAACCTGTTTAGTCTCCTATGGCTAAAGCTCTTCACGATGCATTTGGTATGCCAAAAAGGTC	856	
Db	577	CAAACCTGTTTAGTCTCCTATGGCTAAAGCTCTTCACGATGCATTTGGTATGCCAAAAAGGTC	636	
QY	857	TTATGACTCAATCCACGCTTTATCTGTGTGACCAATATGATCCCTTGACGGACCAACCCGTG	916	

[illegible]

Qy	1157	CTGAGAGATCCAATTTGTTTCTTTTCAGATATCGTAGCGTGTCATACGGTTCAATGTTTGACG	12116
Db	935	CTGAAGATCCAATCGTTTCTTTTCAGATATCGTAGCGGTATCATACGGTTCAATGTTTGACG	994
Qy	1217	CAACTCAAACTTAAAGTTATGGAAGTTTCACGGATCACAAATTGGTTTAAAGTTGTATCATGGT	1276
Db	995	CAACTCAAACTTAAAGTAATGGAAGTTGACGGATCACAAATTGGTTTAAAGTTGTATCATGGT	1054
Qy	1277	ATGCAATGAAATGTCTTTACACTGCTCAAACTTGTTCGTACACTTGGATATTTTGCAAAAA	1336
Db	1055	ATGACAACGAATGTCTTTACACTGCTCAAACTTGTACGTACTCTTTGAGTACTTCGCAAAAA	1114
Qy	1337	TCGCTAAATAA	1347
Db	1115	TTGCTAAATAA	1125

RESULT	13
LOCUS	AF520813
DEFINITION	Streptococcus pyogenes nephritis-associated plasmin receptor (naplr) gene, complete cds.
ACCESSION	AF520813
VERSION	AF520813.1 GI:21686541
KEYWORDS	.
SOURCE	Streptococcus pyogenes
ORGANISM	Streptococcus pyogenes Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
REFERENCE	1 (bases 1 to 2326) Yoshizawa,N., Yamakami,K. and Fujino,M. Nephritis-associated plasmin receptor (naplr) of group A Streptococci pyogenes T type 4 Unpublished
JOURNAL	2 (bases 1 to 2326) Yoshizawa,N., Yamakami,K. and Fujino,M. Direct Submission
AUTHORS	Submitted (07-JUN-2002) Public Health, National Defense Medical College, 3-2, Nakamichi, Tokorozawa, Saitama 359-8513, Japan
TITLE	
JOURNAL	

ORIGIN

[illegible]

SOURCE		Streptococcus pyogenes M1 GAS
ORGANISM		Streptococcus pyogenes M1 GAS Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
REFERENCE		1 (bases 1 to 10053)
AUTHORS		Ferretti,J.J., McShan,W.M., Adjic,D., Savic,D., Savic,G., Lyon,K., Primeaux,C., Sezate,S.S., Surorov,A.N., Kenton,S., Lai,H., Lin,S., Qian,Y., Jia,H.G., Najaf,F.Z., Ren,Q., Zhu,H., Song,L., White,J., Yuan,X., Clifton,S.W., Roe,B.A. and McLaughlin,R.E. Complete genome sequence of an M1 strain of Streptococcus pyogenes Proc. Natl. Acad. Sci. U.S.A. 98 (8), 4658-4663 (2001)
TITLE		21192684
MEDLINE		PUBMED
REFERENCE		2 (bases 1 to 10053)
AUTHORS		Ferretti,J.J., McShan,W.M., Adjic,D., Savic,D., Savic,G., Lyon,K., Primeaux,C., Sezate,S.S., Surorov,A.N., Kenton,S., Lai,H., Lin,S., Qian,Y., Jia,H.G., Najaf,F.Z., Ren,Q., Zhu,H., Song,L., White,J., Yuan,X., Clifton,S.W., Roe,B.A. and McLaughlin,R.E. Direct Submission Submitted (10-Apr-2001) Department of Microbiology and Immunology, University of Oklahoma Health Sciences Center, 940 SL Young Blvd, Oklahoma city, OK 73104, USA Location/Qualifiers 1. .10053 /organism="Streptococcus pyogenes M1 GAS" /mol_type="genomic DNA" /strain="SF370" /serotype="M1" /db_xref="taxon:160490" 205. .1215 /gene="plr" /note="synonym: SPY0274" 205. .1215 /gene="plr" /EC_number="1.2.1.12" /note="The N-terminal amino acid sequence of this ORF has been determined from a spot isolated by 2-D gel electrophoresis from another strain of S. pyogenes. Hogan, D. A., Whittion, M. M., Rogers, J. and R. A. VanBogelen. 2000. Two-dimensional gel electrophoresis map of Streptococcus pyogenes proteins. Unpublished data.; Best Blastp hit = sp P50457 G3P_STRPY_GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE [GAPDH] (PLASMINOGEN-BINDING PROTEIN) (PLASMIN RECEPTOR) >gi 285385 pir A42963 glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - Streptococcus sp >gi 153763 gb AA26953.1 (M95569) plasmin receptor [Streptococcus pyogenes]" /codon_start=1 /transl_table=11 /product="glyceraldehyde-3-phosphate dehydrogenase, plasmin receptor" /protein_id="AAK33348.1" /db_xref="GI:13621550"
gene		/translation="MVVKVINGFGRIGRIFLAFRIONIEGVETRLNDTPNMLAHL LKYDTQGFPGTVEKGEPEUNGFIKVSARDPENIDMDAGDVEIVLEATGFPK KEAKKHANGIKAKXVITAPGDNDXTVPNTNHILDGTETVISGASTCNCLAPM AKALHDAGIOKLMTIHTAYTGQMILDPGHRRGLDARRAGAANIYSTNGAAKAI PIVLPENLKDGAARVPVTGSVELTVLTDKNVSDENIAAKKAANSDFGYTED PLTSSDIVGVSIGSLFDATQTKNVENDGSQLVKVSVYDNMSYTAQLVRTLEYFAKI AK"
CDS		/complement(1699..2439) /gene="SPY0276" Complement(1699..2439) /gene="SPY0276"
gene		/note="Best Blastp hit = pir H69334 glutamine transport protein glnQ - Archaeoglobus fulgidus >gi 2649950 gb AB90561.1 (AE001058) glutamine ABC transporter, ATP-binding protein (glnQ) [Archaeoglobus fulgidus]" /codon_start=1 /transl_table=11
CDS		/product="putative amino acid ABC transporter (ATP-binding protein)"
FEATURES		source
gene		/protein_id="AAK33349.1" /db_xref="GI:13621551" /translations="MNSIIIEIKNLKKSYGSNEVLKDISLVNKGEVSIIGSGSGS IVALRSINLLDEPSAQIILFHGBDVLAEHYNLTTHREKLMGVFQSNLFENLNVLNA STLRSTVLKRDAQAQIAKENLNAVGMTQYWAQKPOLSGQKORVAIALSVNP EAMLFBPTSAIDPEMVGEVLKTMOOLAKSLGLTMIIIVTHEMEFARDVSDRIIFMDKGL ITREGSQOIIFENFTQDRKEFLQRFLK" complement(2432..4000) /gene="SPY0277" complement(2432..4000) /gene="SPY0277" /note="Best Blastp hit = pir S77250 hypothetical protein - Synechocystis sp. (strain PCC 6803) >gi 1852264 dbj BAAL7894.1 (D90907) glutamine-binding periplasmic protein [Synechocystis sp.]" /codon_start=1 /transl_table=11 /product="putative glutamine-binding periplasmic protein" /protein_id="AAK33350.1" /db_xref="GI:13621552" /translations="MKKILSCILALLFGMSRAQNYLRVGMEAAYAFENNQTQ DASNGAVPIEGTSQYANGYDVQAKVAKAMKELLVVKTSWTGLIPALTSCSKIMDIA AGMPTKERNEISFSNSSYTSQPVLVVTTANGKYADATLSLKDFSGAKVTAQQGVHVHN LLTLQKGAKIQTFMGDFSQMRQALTSVDIVAYSERPEAMTAEDASRUMITLLKGGF AVAESDAATAVGNMKDNDMATVNQLEGFSQTRDMALMDMDVTPQVKKAEADAKAS FLQGMWALFKGNWKQFLRGCTMTLLISMVGTITGLFILLIGIFRTAPKARKVAALG OKLFGWLLTIYIIFRGTPMWIQSMVIYYGTACAFGISIDRTLAAIFIYSINTGAYMS EIVRGGIFAVDKGQKFAATALGPTHGQTRKRVLPQVRNLIIPATNGEVEVINIKDTSV LNVISVVELYFSGNTVATQTYQYQFTTIIAIFYVLTFTVTRLRYIERREDADTYT TGANQMQIAEVSNV" 4198. .6096 /genes="SPY0278" 4198. .6096 /genes="SPY0278" /note="Best Blastp hit = no blast hits" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAK33351.1" /db_xref="GI:13621553" /translations="MKKILMTLVLCFSLIGIRIKAADVYSITNYEGOLLISKENTAR FEKVITYQEDTSYNGQYISLGRTGHPAGFAIDOKPKVEYONGQVQPVSOEFSDLG GYRLKYNAGQDKVDVYVITQWLHLLTAYQDVAELNWTPISDMDKLEKVLTVTFT FTDIQDNSWAHRGGYQKKPQVLKQEGNSRYQINAKVNSQGLEHWKALKKKEPV DVSTSKNKIVALETKISRRTLLQLLFGKVIPLVEGVFLMLQLOFTRLKKQFNRYH LANHTSHSEVEDLSPLVLTQAIYQCSFAYLSPASQSKLLIPKGYTFEALVQATL LDLDOKVLLLTKEGKAYLEISOLDRTVDEEAFLDMAFGNKVTLPVDQLFSQHYD ADTIKOLKTTYKGLKEQSVROSSQVVKAMEKASAATINNVLTKLINLPDYROM TPARKKSNSVQSLGCLLLILNSGLLIYAKESGALIYALMLVLTWCLGFYISLKL DQYKUGIETPEGGVRLHQWFQKMIRIDIKFEDVUAIEGLVWNRVLYATLECYAK KVERYLKVRHIALPEYVQVRPCELSMNMYATTPTFVSSLSATTSNFVSMSGGGS CGGGGGAGF" /genes="bacA" /notes="synonym: SPY0280" 6162. .7001 /genes="bacA" 6162. .7001 /genes="bacA" /note="Best Blastp hit = gb AAD50462.1 AF169967_5 (Af169967) BacA [Flavobacterium johnsoniae]" /codon_start=1 /transl_table=11 /product="putative undecaprenol kinase (bacitracin resistance protein)" /protein_id="AAK33352.1" /db_xref="GI:13621554" /translations="MLUIELKXAIFFGIIEGTIEWL FVSSSTGHLLIVOEIRLNQDKA FIDMFNIIVGLGAIIVAMLIYFERLNRPFGQTRAEVOLTWQLWKLVIACIPSILIA VPDNMFEAHFYFMVPIETALIVYGIATFWIEKRNQAQEPVATLARMYSYKTAFFIG FOYLSIVPGTSSRSATILGAILIGTSRTVAADTFELAIPTMFGYSGLKAVKFIDLGH HLDFAQVILLILVASLTAFTAVVUSLLAIRFLTDDYVKKHDFTFGKYRIVLGSLLLIYSFFK FVF"
CDS		/note="meca"
gene		7147. .7908


```

/organism="Streptococcus pyogenes MGAS315"
/mol_type="genomic DNA"
/strain="MGAS315"
/serotype="M3"
/db_xref="taxon:198466"
/notes="group: A"
301..660
/gene="rnpA"
/notes="synonym: SpyM3_0175"
301..660
/gene="rnpA"
/notes="best blastp match gb|AAK33323.1| (AE006491) putative ribonuclease P protein component [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative ribonuclease P protein component"
/protein_id="AAW78782.1"
/db_xref="GI:21903895"
/transl_table=11
/translation="MKTYRVKREKDFQALFKDKSTANRKFVYHLNKGQDHFVGI SVGKKIGNAVTRNAVKRKIRHVIMALGHQKLSDFVVIARKGVESLEYQLQNLHHV LKAQLLEKGFSEERK"
626..1453
/gene="SpyM3_0176"
626..1453
/gene="SpyM3_0176"
/notes="best Non-GAS blastp hit: gb|AAK76106.1| (AE007493) SpolIIV family protein [Streptococcus pneumoniae TIGR4]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAW78783.1"
/db_xref="GI:21903896"
/transl_table=11
/translation="MKVLRVKKNIKIARIVPLVLLVACRGVEVTAQSSGWDOLVY LPARQWLSFDGSIQGIITLTLRLMPLFNQIKSSQKQMDIQELBELQKRY AGDTQTRMLAESGALYKGVNYPYSLPILQMPVIALFQALTRVSKGTGF LWVLAQHLXLYLPVLAAYFTLSTLNLAAKKNVMTVMVYMLPMIFMGFTL ASGVLYTWTSNAFQVQVQLLLNPPFKIIAERQRLANEERKERRARKKARKK"
1465..2379
/gene="SpyM3_0177"
1465..2379
/gene="SpyM3_0177"
/notes="best Non-GAS blastp hit: gb|AAK76105.1| (AE007493) jag protein, putative [Streptococcus pneumoniae TIGR4], and gb|AAL00654.1| (AE008549) hypothetical protein [Streptococcus pneumoniae R6]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAW78784.1"
/db_xref="GI:21903897"
/transl_table=11
/translation="MVLFTGVVEAIEITGLQELGLPRLKAHKVISEKKGFLQFGK KPAQVDEIGSDTKVYKADKATRGVPEDINQNAFVNSADVEPEEKATQRLAEAD TKVYPLMSDSQAIPSNLAETVETKAPQSPVPESEVPQDAGDGFQIEKAAQ EVSDYVTKIYEMDIEATVETSNRRQINLOIETPAGRVIGYHGKVLKSLQLAQN LFDRYSKNFSLNVDYVHEHRTETLIDFTQKAVKRVLSGQDYMDSNSERKIVH KTVSSIEGVDSVSEGNDRYVYVSLQR"
2694..2828
/gene="rpnH"
/notes="synonym: SpyM3_0178"
2694..2828
/gene="rpnH"
/notes="best blastp hit: gb|AAK33326.1| (AE006491) 50S ribosomal protein L34 [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="50S ribosomal protein L34"
/protein_id="AAW78785.1"
/db_xref="GI:21903898"
/transl_table=11
/translation="MKRTYQPSKIRQRKHGFRHMTGNGRRVLAARRKGRKVLSA"
3102..3806
/gene="SpyM3_0179"
3102..3806
/gene="SpyM3_0179"

```

```

3102..3806
/gene="SpyM3_0179"
/notes="best Blastp hit: gb|AAK33327.1| (AE006492) putative N-acetylmannosamine-6-P epimerase [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative N-acetylmannosamine-6-P"
/protein_id="AAW78786.1"
/db_xref="GI:21903899"
/transl_table=11
/translation="MPDKPTKLMQKGLGIIIVSCQALPCEPLYSETGGIMPLLAKA AQAGAVCIRANSVRDKEIOAJTDPLIGIIVKDYPOEPFITATMAVVDOLALANI AVIAMDCKRDRHGLDIAIFIRQVKYFNQLLMADISTDEGLVHAQAGIDFVGTT LSGTIPYRQAGPDVALIEALCKRAGIAVIAEGKIHSPFEAKKINDLGLVAGIVVGGAI TRPKIEAERFIEALKS"
3855..4913
/gene="SpyM3_0180"
3855..4913
/gene="SpyM3_0180"
/notes="best Blastp hit: gb|AAK33328.1| (AE006492) putative sugar transporter sugar binding lipoprotein [Streptococcus pyogenes M1 GAS]; possible pseudogene aligns with N-terminal portion (amino acids 1-352)"
/codon_start=1
/transl_table=11
/product="putative sugar transporter sugar binding lipoprotein"
/protein_id="AAW78787.1"
/db_xref="GI:21903900"
/transl_table=11
/translation="MNMKGLASVLMGASVLGAACGKSKOKERAGSKSDTAKTEITW WAFVFTQERAEADGVTEYKELIAAFKKEKPEIKVNDKLIQASGADTAYMIPSSA PVDLFDAPGRIIOYKNGKGLADNLDFKTEFTKVDNDKLIQASGADTAYMIPSSA PFYALNKKFKDAGVLDLVKEGWTDDPEKVLAKDKGYNPGSFFANGGGQGGPR AFFANLSSHITDDKVTYTDANSIKAMTKISNWKDGLMNGSQYDSDGADIONFA NGQTSFILNAPAQFGIQAALLEASKVDYLEIPFSDGKPELELYLVNGFAVFNKOE QKVAASKTFQFIADDEK"
4989..5174
/gene="SpyM3_0181"
4989..5174
/gene="SpyM3_0181"
/notes="best Blastp hit: gb|AAK33328.1| (AE006492) putative sugar transporter sugar binding lipoprotein [Streptococcus pyogenes M1 GAS]; possible pseudogene aligns with C-terminal portion (amino acids 379-439)"
/codon_start=1
/transl_table=11
/product="putative sugar transporter sugar binding lipoprotein"
/protein_id="AAW78788.1"
/db_xref="GI:21903901"
/transl_table=11
/translation="MEKIAETKFPSPYNTIDGFAEMRTLWFPVQALNSGDEKPED ALKAFTEKANKTKTKTQ"
5299..6165
/gene="SpyM3_0182"
5299..6165
/gene="SpyM3_0182"
/notes="best Blastp hit: gb|AAK33329.1| (AE006492) sugar transport system (permease) [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative sugar transport system (permease)"
/protein_id="AAW78789.1"
/db_xref="GI:21903902"
/transl_table=11
/translation="MRETLISYAFAPLVVFFVILPIMMGFVTSFFNYSMTDFTF FFLPVVTVGSAVTVKWKIYDPMGILNVLKVAHVIEQINISWLGKHWALLAIIL LITSVGQPIILYIAAMGNIDNSLMEARVDGATEFOFVWNIKWPSLLPTLYLAVITT INSQGFALQILTSGGPNYSTLTMLVLYEKAFKLESEYANTMGVFLAVMTAIIISF AQKILGNDEY"
6178..7008
/gene="SpyM3_0183"
6178..7008
/gene="SpyM3_0183"

```



```

/gene="SpyM3_0183"
/note="best Blastp hit: gb|AAK33330.1| (AE006492) probable
sugar transport protein (permease) [Streptococcus pyogenes
M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative sugar transport protein (permease)"
/protein_id="AM78790.1"
/db_xref="GI:21903903"
/translation="MKKKKLARSVDLTVILCILITLIPFPYHIMTCAKQADTIM
IPQWPKPAPIENFKALVIONPALKMLWNSVFIISVTMLFLVCGTSSLAGYALAKRF
YQRLIFSIPTAAVALPKQVVLPLVIRVNFVMDHLAAVILVGVMPFVLMKQF
SENIPTELESAKIDCGCEIRTFNVAFPIVKPGFAALAIFFINWDFMQLVMLT
SRENLTISLGVAWQAEWATNYGLIMAGAAVAPEIVTVFLVFKSTFQTGIMCAVKG
"
7165..7827
/gene="SpyM3_0184"
/contig="7827
/gene="SpyM3_0184"
/note="best non-GAS blastp hit: gb|AAK75756.1| (AB007460)
hypothetical protein [Streptococcus pneumoniae TIGR4]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"

Query Match      58.0%; Score 780.6; DB 1; Length 50029;
Best Local Similarity 94.8%; Pred. NO. 2.4e-147;
Matches 807; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 497 TCGATGGTACTGTTGAAGCTTAAGAGCTGATCGAAGTTAAGCGTCAATTTCGTTAAAG 556
Db 25967 TTGATGGAAACAGTTGAAGTTAAAGAGGTGATTTGAAGTAAACGGAACTTCTCAAG 26026

QY 557 TTTCGTCTGAACGCGAACCAGCAACATTTGATGGGCTTACTGATGGCGTAGAAATCGTTC 616
Db 26027 TTTCGTCTGAACGTGATCCAGAAACATCGATGGCAACTGATGGGCTTGAATCGTTC 26086

QY 617 TTGAAGCAACTAGTTCTTTCTGCTAAAGAAAGAGCTGCTGAAACACACTTACATGCTAACG 676
Db 26087 TTGAAGCAACTGGTTCTTTCTGCTAAAGAAAGAGAGCTGAAACACACTTACATGCTAACG 26146

QY 677 GTGCTAAAAAAGTTGTTATACAGCTCTCTGCTGGAACAGCGTTAAACAGTTGTTTCA 736
Db 26147 GTGCTAAAAAAGTTGTTATACAGCTCTCTGCTGGAACAGCGTTAAACAGTTGTTTCA 26206

QY 737 ACACTAACACGACATTTCTGACGCTACTGAAACAGTTATCTCAGGTGCTTTCACTACTA 796
Db 26207 ACACTAACACGACATTTCTGACGCTACTGAAACAGTTATCTCAGGTGCTTTCACTACTA 26266

QY 797 CAACCTGTTAGCTCTATGCTTAAGCTCTTACAGTGGATTTGGTATCCAAAAGGTC 856
Db 26267 CAACCTGTTAGCTCTATGCTTAAGCTCTTACAGTGGATTTGGTATCCAAAAGGTC 26326

QY 857 TTATGACTACAAATCCACGCTTATCTGCTGACCAAAATGATCCTTTGACGGACCAACCGTG 916
Db 26327 TTATGACTACAAATCCACGCTTATCTGCTGACCAAAATGATCCTTTGACGGACCAACCGTG 26386

QY 917 GTGCTGACCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 26387 GTGCTGACCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 26446

QY 977 CTGCTAAAAGCTATCGGTCTTGTATCCAGAAATGAATGGTAAACTTTGATGGTCTGCAC 1036
Db 26447 CTGCTAAAAGCTATCGGTCTTGTATCCAGAAATGAATGGTAAACTTTGATGGTCTGCAC 26506

QY 1037 AACGTGTTCTCTGCTCCAACTGATCAGTAACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1096
Db 26507 AACGTGTTCTCTGCTCCAACTGATCAGTAACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 26566

QY 1097 TTTCGTGTTGACGAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTGGTGGTGGTGGT 1156
Db 26567 TTTCGTGTTGACGAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTGGTGGTGGTGGT 26626

```

```

QY 1157 CTGAAGATCCAAATGTTTCTTTCAGATATCGTAGCGGTGCATACGGTTCATTTGTTGACG 1216
Db 26627 CTGAAGATCCAAATCGTTTCTTTCAGATATCGTAGCGGTGCATACGGTTCATTTGTTGACG 26686

QY 1217 CAACCTCAAACTAAAGTTATGGAAGTTTGACGGATCACAATTTGGTTAAAGTTGTATCATGGT 1276
Db 26687 CAACCTCAAACTAAAGTTAATGGAAGTTTGACGGATCACAATTTGGTTAAAGTTGTATCATGGT 26746

QY 1277 ATGACAATGAAATGTTCTTACACTGCTCAACTTGTTCGTACACTTGTAGTATTTTGCAAAA 1336
Db 26747 ATGACAACGAAATGTTCTTACACTGCTCAACTTGTTCGTACACTTGTAGTATTTTGCAAAA 26806

QY 1337 TCGCTAAATAA 1347
Db 26807 TTGCTAAATAA 26817

```

Search completed: March 31, 2004, 00:58:07
Job time : 5274 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 18:05:29 ; Search time 566 Seconds
(without alignments)
10110.111 Million cell updates/sec

Title: US-10-650-369-21

Perfect score: 1347

Sequence: 1 atgaaaaataacaggat.....ttgcaaaaatcgctaaataa 1347

Scoring table: IDENTITY NUC

Gapop 10_0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseqn 29Jan04: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002s: *
7: geneseqn2003as: *
8: geneseqn2003bs: *
9: geneseqn2003cs: *
10: geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1347	100.0	1347	6	ABA91327 Streptococcus
2	815.8	60.6	1011	6	ABA91328 Streptococcus
3	815.8	60.6	1011	6	ABA91248 Streptococcus
4	780.6	58.0	1011	7	ACA50325 Streptococcus
5	780.6	58.0	1125	2	AAQ70705 Plasmin r
6	780.6	58.0	1125	4	AAC66140 Streptococcus
7	777.6	57.7	1008	6	ABN70591 Streptococcus
8	777.6	57.7	1035	6	ABN69737 Streptococcus
9	702.2	52.1	1011	6	ABA91329 Streptococcus
10	702.2	52.1	1011	6	ABA91249 Streptococcus
11	699	51.9	110000	6	Continuation (18 o
12	696	51.7	1008	6	ABN71527-17
13	683	50.7	1080	4	AAS55858 Streptococcus
14	678.2	50.3	1011	6	ABA91332 Streptococcus
15	678.2	50.3	1011	6	ABA91252 Streptococcus
16	678.2	50.3	1080	4	AAS55434 Streptococcus
17	678.2	50.3	1080	7	ACA50132 Streptococcus
18	678.2	50.3	7571	2	AAV52135 Streptococcus
19	678.2	50.3	110000	7	ABSS56454-18
20	676.6	50.2	1011	6	ABA91330 Streptococcus
21	676.6	50.2	1011	6	ABA91250 Streptococcus
22	675.2	50.1	1005	7	ABX07806 Streptococcus
23	660.2	49.0	1000	2	AAV27350 Streptococcus

24 660.2 49.0 1000 6 ABQ84818 Abq84818 S. pneumo
25 660.2 49.0 1000 9 ADC45134 Adc45134 S. pneumo
26 640.8 47.6 5448 2 AAV65229 Aav65229 DNA encod
27 639.8 47.5 1011 6 ABA91331 Aba91331 Streptoco
28 639.8 47.5 1011 6 ABA91251 Aba91251 Streptoco
29 620.6 46.1 1011 7 ACA48200 Aca48200 Prokaryot
30 567.8 42.2 65589 6 ABA90521-23 Continuation (24 o
31 495.6 36.8 1002 9 ADC91488 Adc91488 E. faeciu
32 481.4 35.7 110000 6 ABA90521-05 Continuation (6 of
33 479.6 35.6 1002 4 AAS53113 Aas53113 Enterococ
34 478 35.5 999 7 ACA33240 Aca33240 Prokaryot
35 478 35.5 4023 2 AAX13023 Aax13023 Enterococ
36 478 35.5 4023 6 ABS98818 Abs98818 Enterococ
37 454 33.7 110000 6 ABQ67196-1 Continuation (2 of
38 454 33.7 110000 6 ABQ69245-25 Continuation (26 o
39 453.6 33.7 1005 7 ACA27465 Aca27465 Prokaryot
40 450.8 33.5 1011 7 ACA36438 Aca36438 Prokaryot
41 450.8 33.5 110000 6 ABA03041-25 Continuation (26 o
42 447.6 33.2 4230 6 ABQ70965 Abq70965 Listeria
43 426.8 31.7 1011 4 AAS33908 Aah53908 S. epider
44 426.8 31.7 1047 6 ABN93213 Abn93213 Staphyloc
45 426.8 31.7 4642 4 AAN54726 Aan54726 S. epider

ALIGNMENTS

RESULT 1
ABA91327
ID ABA91327 standard; DNA; 1347 BP.

XX ABA91327;
XX 29-AUG-2003 (revised)
DT 08-APR-2002 (first entry)
XX Streptococcus GapC multiepitope fusion PolyGap4 DNA.
XX PolyGap4; GapC; plasmin binding protein; epitope; infection; vaccine;
XX immunisation; mastitis; therapy; gene; ds.
XX Streptococcus dysgalactiae.
OS Streptococcus agalactiae.
OS Streptococcus parauberis.
OS Chimeric.
XX WO200196379-A2.
XX 20-DEC-2001.
XX 11-JUN-2001; 2001WO-CA000836.
XX 12-JUN-2000; 2000US-0211247P.
XX (UYSA-) UNIV SASKATCHEWAN.
XX Potter AA, Perez-Casal J, Fontaine M;
XX WPI; 2002-098051/13.
XX P-P8DB; AAM50664.
XX Novel GapC multiple epitope fusion polypeptide comprising antigenic
XX determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
XX parauberis, or S. iniae GapC protein useful for treating mastitis in
XX vertebrates.
XX Example 4; Fig 6A-C; 116pp; English.
XX The present sequence is that of a chimeric gene encoding a multiple
XX epitope fusion protein (see AAM50664), designated PolyGap4, comprising
XX the entire amino acid sequence of the Streptococcus dysgalactiae GapC
XX plasmin binding protein in addition to unique amino acid sequences from
XX the Streptococcus parauberis and Streptococcus agalactiae GapC proteins.

The gapC gene sequences used to construct the chimeric gene were prepared by PCR amplification of selected polynucleotide sequences from the genomic gapC genes using the primers given in AB91335-42. After assembly, the chimeric gene was inserted into plasmid pAA555. PolyGala is an example of novel GapC multiple epitope fusion proteins of the invention that comprise epitopes from 1 or more of *S. dysgalactiae*, *S. agalactiae*, *S. paratyphi*, *Streptococcus uberis* and *Streptococcus* *iniae* (see AA50665-69). The multiple epitope fusion proteins are produced using claimed host-vector systems and used in claimed vaccines for treating or preventing a bacterial infection in a vertebrate, especially a streptococcal infection, and particularly mastitis. They are also used in claimed methods of detecting *Streptococcus* antibodies. The multiple epitope protein is capable of eliciting broad immunity against a variety of streptococcal infections while minimising the number of antigens present in the final formulation and concomitantly reducing production costs. (Updated on 29-AUG-2003 to standardise OS field)

Sequence 1347 BP: 401 A: 251 C: 297 G: 398 T: 0 U: 0 Other: 0

Query Match	100.0%	Score 1347;	DB 6;	Length 1347;
Best Local Similarity	100.0%;	Fred. No. 0;		
Matches 1347;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAAAAAATAACAGGATATTTTATATGCTTCTTGCAGTCAATATTTCTGTCGATGC	60	
Db	1	ATGAAAAAATAACAGGATATTTTATATGCTTCTTGCAGTCAATATTTCTGTCGATGC	60	
QY	61	CAGGCAAACTACGGATCCGGTATCGTAGTTTAAAGTTGGTATTAAACGGTTTCGGTTCGTATC	120	
Db	61	CAGGCAAACTACGGATCCGGTATCGTAGTTTAAAGTTGGTATTAAACGGTTTCGGTTCGTATC	120	
QY	121	GSAGCTCTTGCAATCCCGTCGTAATCAAAATGTTGAGAGTGTTCGAAGTAACCTCGTATCAAC	180	
Db	121	GSAGCTCTTGCAATCCCGTCGTAATCAAAATGTTGAGAGTGTTCGAAGTAACCTCGTATCAAC	180	
QY	181	GACCTTACAGATCCAAAACATGCTTGCACACTTGTGAAATACGATACAACTCAAGGACGT	240	
Db	181	GACCTTACAGATCCAAAACATGCTTGCACACTTGTGAAATACGATACAACTCAAGGACGT	240	
QY	241	TTTGACGGAACTGTTGAAAGTTAAAGAGTGGATTTGAAGTAAACGGAAACCTTCATCAAA	300	
Db	241	TTTGACGGAACTGTTGAAAGTTAAAGAGTGGATTTGAAGTAAACGGAAACCTTCATCAAA	300	
QY	301	GTTTCTGCTGAACGCGTATCCAGAAAAACATCGACTGGGCAACTGACGGTGTTCGAAATCGTT	360	
Db	301	GTTTCTGCTGAACGCGTATCCAGAAAAACATCGACTGGGCAACTGACGGTGTTCGAAATCGTT	360	
QY	361	CTGGAAGCACTCGAGGGTACTGTAGAAGTTTAAAGATGGTGGATTTGACGTTTAAACGGAAAA	420	
Db	361	CTGGAAGCACTCGAGGGTACTGTAGAAGTTTAAAGATGGTGGATTTGACGTTTAAACGGAAAA	420	
QY	421	TTTCATTAAAGTTTCTGCTGAAAAAGATCCAGAACAAATTCGACTGGGCAACTGACGGTGT	480	
Db	421	TTTCATTAAAGTTTCTGCTGAAAAAGATCCAGAACAAATTCGACTGGGCAACTGACGGTGT	480	
QY	481	GAAATCGTTTCTTGAATCGATGGTACTGTTGAAAGTTTAAAGAGGTTGATTCGAAAGTTAAC	540	
Db	481	GAAATCGTTTCTTGAATCGATGGTACTGTTGAAAGTTTAAAGAGGTTGATTCGAAAGTTAAC	540	
QY	541	GGTCAATTTGTTAAAGTTTCTGCTGAACGGCAACAGCAAACTTGACTGGGCTACTCAT	600	
Db	541	GGTCAATTTGTTAAAGTTTCTGCTGAACGGCAACAGCAAACTTGACTGGGCTACTCAT	600	
QY	601	GGCGTAGAATCGTTTCTTGAAGCAACTAGTTTCTTCTTAAAAAAGAAAGCTGCTGAAAAA	660	
Db	601	GGCGTAGAATCGTTTCTTGAAGCAACTAGTTTCTTCTTAAAAAAGAAAGCTGCTGAAAAA	660	
QY	661	CACCTTACATGCTAAACGGTGTCTAAAAAGTTGTTATCAACGCTCCTGGTGGAAACGACGTT	720	
Db	661	CACCTTACATGCTAAACGGTGTCTAAAAAGTTGTTATCAACGCTCCTGGTGGAAACGACGTT	720	
QY	721	AAAAACAGTTGTTTCTCAACACTAAACACGACATTCCTTCGACGGTACTCGAAACAGTTATCTCA	780	

PT Novel GapC multiple epitope fusion polypeptide comprising antigenic
 PT determinant of *Streptococcus dysgalactiae*, *S. agalactiae*, *S. uberis*, *S.*
 PT *parvauberis*, or *S. iniae* GapC protein useful for treating mastitis in
 PT vertebrates.
 XX
 PS Example 3; Fig 1A-B; 116pp; English.

XX The present sequence is that of the coding region of the gene encoding
 CC the GapC plasmid binding protein, DysgapC (see AAM50665), of
 CC *Streptococcus dysgalactiae* ATCC 43078, an isolate from a case of bovine
 CC mastitis. The gene was obtained by PCR amplification of chromosomal DNA
 CC using the primers given in ABA91248-34. The invention relates to novel
 CC GapC multiple epitope fusion proteins that comprise epitopes from 1 or
 CC more of *Streptococcus dysgalactiae*, *Streptococcus agalactiae*,
 CC *Streptococcus parvauberis*, *Streptococcus uberis* and *Streptococcus iniae*
 CC (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664). The
 CC multiple epitope fusion proteins are produced using claimed host-vector
 CC systems and used in claimed vaccines for treating or preventing a
 CC bacterial infection in a vertebrate, especially a streptococcal
 CC infection, and particularly mastitis. They are also used in claimed
 CC methods of detecting *Streptococcus* antibodies. The multiple epitope
 CC proteins are capable of eliciting broad immunity against a variety of
 CC streptococcal infections while minimising the number of antigens present
 CC in the final formulation and concomitantly reducing production costs
 XX
 SQ Sequence 1011 BP; 295 A; 200 C; 216 G; 300 T; 0 U; 0 Other;

Query Match 60.6%; Score 815.8; DB 6; Length 1011;
 Best Local Similarity 97.4%; Pred. No. 3.4e-207;
 Matches 829; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
 QY 497 TCGATGGTACTCTGAAGTTAAGAGGTGGATTTCGAGTTAAGCGTCAATTTGTTAAG 556
 DB 161 TTGACGGAACTGTGTAAGTTAAGAGGTGGATTTCGAGTTAAGCGTCAATTTGTTAAG 220
 QY 557 TTTCTGCTGAACGCAACAGCAACATTTGACTGGGCTACTGATGGCGTAGAAATCGTTTC 616
 DB 221 TTTCTGCTGAACGCTGATCCAGAAACATCGACTGGGCAACTGACGGTGTGAATCGTTTC 280
 QY 617 TTGAAGCACTAGTTTCTTTGCTTAAAGAGAGCTGCTGAAACACACTTACATGCTAACG 676
 DB 281 TCGAAGCACTAGTTTCTTTGCTTAAAGAGAGCTGCTGAAACACACTTACATGCTAACG 340
 QY 677 GTGCTAAAAAGTTGTTATCACAGTCTCTGGTGGAAACGACGTTAAACAGTTGTTTCA 736
 DB 341 GTGCTAAAAAGTTGTTATCACAGTCTCTGGTGGAAACGACGTTAAACAGTTGTTTCA 400
 QY 737 ACACTAACCAACAGCACTTTGACGCTGCTGAAACAGTTATCTCAGGTGCTTCATGCTACTA 796
 DB 401 ACACTAACCAACAGCACTTTGACGCTGCTGAAACAGTTATCTCAGGTGCTTCATGCTACTA 460
 QY 797 CAAACTGTTTACGCTCTATGCTTAAAGCTCTTTCAGATGCTATTCGATGCCAAAGGTC 856
 DB 461 CAAACTGTTTACGCTCTATGCTTAAAGCTCTTTCAGATGCTATTCGATGCCAAAGGTC 520
 QY 857 TTATGACTACATCCACGCTTATCTATGCTGACCAATGATCTTCACGAGCAACACCGTG 916
 DB 521 TTATGACTACATCCACGCTTATCTATGCTGACCAATGATCTTCACGAGCAACACCGTG 580
 QY 917 GTGGTGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
 DB 581 GTGGTGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
 QY 977 CTGCTAAAGCTATCGGCTCTTGTATCCAGAAATGGAATGGTAAATGATGCTGCTGCTGCTG 1036
 DB 641 CTGCTAAAGCTATCGGCTCTTGTATCCAGAAATGGAATGGTAAATGATGCTGCTGCTGCTG 700
 QY 1037 AACGTTTCTGTTTCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1096
 DB 701 AACGTTTCTGTTTCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 760
 QY 1097 TTTCTGTTTCAAGAAATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1156

DB 761 TTTCTGTTTCAAGAAATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 820
 QY 1157 CTGAGATCCCAATTTGTTTCTTCAGATATCTGAGCGTGTCTATCGGTTTCATGTTTTCACG 1216
 DB 821 CTGAGATCCCAATTTGTTTCTTCAGATATCTGAGCGTGTCTATCGGTTTCATGTTTTCACG 880
 QY 1217 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCAACAATTTGGTTAAAGTTGTTATCATGGT 1276
 DB 881 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCAACAATTTGGTTAAAGTTGTTATCATGGT 940
 QY 1277 ATGACATGAAATGCTTCTTACACCTGCTCAACTTGTGCTACACTTGAGTATTTTGCACAAA 1336
 DB 941 ATGACATGAAATGCTTCTTACACCTGCTCAACTTGTGCTACACTTGAGTATTTTGCACAAA 1000
 QY 1337 TCGCTAAATAA 1347
 DB 1001 TCGCTAAATAA 1011
 RESULT 3
 ABA91248
 ID ABA91248 standard; DNA; 1011 BP.
 XX
 AC ABA91248;
 XX
 DT 04-APR-2002 (first entry)
 XX
 DE Streptococcus dysgalactiae gapC gene.
 XX
 KW GapC; plasmin-binding protein; infection; mastitis; vaccine; diagnosis;
 XX gene therapy; gene; ds.
 XX
 OS Streptococcus dysgalactiae.
 XX
 FN WO200196381-A2.
 XX
 PD 20-DEC-2001.
 XX
 PP 11-JUN-2001; 2001WO-CA000838.
 XX
 PR 12-JUN-2000; 2000US-0211022P.
 XX
 PA (UYSA-) UNIV SASKATCHEWAN.
 XX
 PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
 XX
 DR WPI; 2002-130725/17.
 XX
 DR P-PSDB; AAM50639.
 XX
 PT Novel isolated GapC protein of *Streptococcus dysgalactiae*, *S. agalactiae*,
 PT *S. uberis*, *S. parvauberis*, or *S. iniae*, useful as vaccine component for
 PT treating streptococcal infection which causes mastitis in vertebrates.
 XX
 PS Example 1; Fig 1A-B; 107pp; English.
 XX
 CC The present sequence is that of the coding region of the GapC plasmid
 CC binding protein gene of *Streptococcus dysgalactiae* ATCC 43078, a clinical
 CC isolate from a case of bovine mastitis. The gene was amplified from
 CC chromosomal DNA by PCR using the primers given in ABA91253 and ABA91254.
 CC The PCR product was cloned into expression vector pET15N, creating
 CC plasmid pET15bgapC. The gene encodes a GapC plasmin binding protein (see
 CC AAM50639) that is capable of eliciting an immune response in a
 CC vertebrate. The invention provides the GapC genes and proteins of 5
 CC *Streptococcus* species, as well as recombinant vectors, host cells and
 CC vaccine compositions comprising GapC polynucleotides or proteins. The
 CC vaccines are used to treat or prevent a bacterial infection, especially a
 CC streptococcal infection, and mastitis in particular (claimed). A
 CC polynucleotide encoding a GapC protein is used in a claimed method of
 CC treating or preventing a bacterial infection, such as a streptococcal
 CC infection, especially mastitis
 XX
 SQ Sequence 1011 BP; 295 A; 200 C; 216 G; 300 T; 0 U; 0 Other;

	Query Match	60.6%; Score 815.8; DB 6; Length 1011;
	Best Local Similarity	97.4%; Pred. No. 3.4e-207;
	Matches 829; Conservative	0; Mismatches 22; Indels 0; Gaps 0;
QY	497	TCGATGTCACGTGTTGAAGTTAAAGAAGCTCGATTCCGAAGTAAACGGTCAATTTGTTAAAG 556
Db	161	TTGACGGAACGTGTTGAAGTTAAGAAGTGGAATTTGAAGTAACGGAACCTTCATCAAAG 220
QY	557	TTTCTGCTGAACCGCAACCGAGCAACAATCACTGGGCTACTGATGGCGTAGAATAATCGTTC 616
Db	221	TTTCTGCTGAACGTGATCCAGAAAACATCACTGGGCAACTGACGGTGTGTGAATAATCGTTC 280
QY	617	TTCAAGCAACACTAGTCTTTTCTTAAAAAGAGAGCTGCTGAAAAACACTTACATGCCTAACG 676
Db	281	TGGAAGCAACACTGTTCTTTTCTTAAAAAGAGAGCTGCTGAAAAACACTTACATGCCTAACG 340
QY	677	GTGCTAAAAAGTTGTTTATCACAGCTCCTCGTGGAAACGAGTTAAACAGTTGTTTTCA 736
Db	341	GTGCTAAAAAGTTGTTTATCACAGCTCCTCGTGGAAACGAGTTAAACAGTTGTTTTCA 400
QY	737	ACACTAACACGACATCTTTGACGGTACTGAACACAGTTATCTCAGGTGCTTCATGTACTA 796
Db	401	ACACTAACACGACATCTTTGACGGTACTGAACACAGTTATCTCAGGTGCTTCATGTACTA 460
QY	797	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTCA CGATGCAATTGTSATCCAAAAAGGTC 856
Db	461	CAAACTGTTTAGCTCCTATGGCTAAAGCTCTTCA CGATGCAATTGTSATCCAAAAAGGTC 520
QY	857	TTATGACTACAATCCAGCCTTATCTGGTGACCAAAATGATCTTGACGGACCAACCGGTG 916
Db	521	TTATGACTACAATCCAGCCTTATCTGGTGACCAAAATGATCTTGACGGACCAACCGGTG 580
QY	917	GTGGTGACCCTTCGTGCTCGTGTGCTGCTGCAAA CATTGTTCCCTAACTCAACTGGTG 976
Db	581	GTGGTGACCCTTCGTGCTCGTGTGCTGCTGCAAA CATTGTTCCCTAACTCAACTGGTG 640
QY	977	CTGCTAAAGCTATCGGTCCTGTTATCCAGAA TTGAATGTATAACTTGATGGTGTGTCAC 1036
Db	641	CTGCTAAAGCTATCGGTCCTGTTATCCAGAA TTGAATGTATAACTTGATGGTGTGTCAC 700
QY	1037	AACGTGTTCTGTTCCAACCTGGATCAGTAACTCAGTTGGTTGTA ACTCTTTGATAAAAAACG 1096
Db	701	AACGTGTTCTGTTCCAACCTGGATCAGTAACTCAGTTGGTTGTA ACTCTTTGATAAAAAACG 760
QY	1097	TTTCTGTGAGGAAATCAACCGTGCTATGAAGCTGCTTCAAACGACAGTTTCGGTTACA 1156
Db	761	TTTCTGTGAGGAAATCAACCGTGCTATGAAGCTGCTTCAAACGACAGTTTCGGTTACA 820
QY	1157	CTGAAGATCCAAATGTTCTTTCAGATATCGTAGCGGTGTCATACGGTTCA TTTGTTGAOC 1216
Db	821	CTGAAGATCCAAATGTTCTTTCAGATATCGTAGCGGTGTCATACGGTTCA TTTGTTGAOC 880
QY	1217	CAACTCAAATAAGTTATGGAAGTTGACGGATCA CAATTTGGTTAAAGTTGTTATCATGGT 1276
Db	881	CAACTCAAATAAGTTATGGAAGTTGACGGATCA CAATTTGGTTAAAGTTGTTATCATGGT 940
QY	1277	ATGACAATGAAATGCTTTACACTGCTCAACT GTTTGTTGTTACACTTGAGTATTTTGC AAAA 1336
Db	941	ATGACAATGAAATGCTTTACACTGCTCAACT GTTTGTTGTTACACTTGAGTATTTTGC AAAA 1000
QY	1337	TCGCTAAATAA 1347
Db	1001	TCGCTAAATAA 1011

RESULT 4

RESUL 4
ACA50325

ACA50325
ID ACA50325 standard; DNA; 1011 BP.

XX
XX
XX

ACAC

2000

DT 19-0

XX

Prokaryotic essential gene #31982.

Antisense; ds; prokaryotic essential gene; cell proliferation; drug design; gene.

Streptococcus pyogenes.

WO200277183-A2.

03-004-2002

21-MAP-2002: 2002WO-IIS009107

21 - MAP-3001 - 2001118 - 0081E042

21-MAR-2001; 2001US-00815242.
06-SEP-2001; 2001US-00948993

06-SEP-2001; 2001US-00948993.
25-OCT-2001: 2001US-0342923D.

25-OCT-2001; 2001US-0342923P;
08-FEB-2002; 2002US-00072851.

08-FEB-2002; 2002US-00072851.
06-MAR-2002; 2002US-03626992.

(E.I.T.T-) E.I.T.T.R.A. PHARM. INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI: 2003-029926/02.

New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 14: SEO ID NO 38195: 1766pp: English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *X. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://www.int/pub/published/pct_sequences

Sequence 1011 BP: 300 A: 206 C: 213 G: 292 T: 0 U: 0 Other:

Query Match 58.0%: Score 780.6: DB 7: Length 1011;

Query Match	58.0%;	score	780.6;	DB	71
Best Local Similarity	94.8%;	Pred.	NQ. 8.5e-198;		

BEST LOCAL SIMILARITY 34.8%;
Matches 807; Conservative
0; Mismatches 44;
FREQ: NO: 8:38-138

497 TCGATGGTACTGTTGAAGTTAAAGAAGGTGGATTCCGAAGTTAACGGTCAATTGTAAAG 556

Qy

QY 857 TTATGACTACAAATCCACGCTTATATCTGTGACCAATGATCTTTGACGGACCAACACCGTG 916
Db 635 TTATGACTACAAATCCACGCTTATATCTGTGACCAATGATCTTTGACGGACCAACACCGTG 694
QY 917 GTGGTGACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 695 GTGGTGACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 754
QY 977 CTGCTAAAGCTATCGGCTTCTGTTATCCAGAAATGAAATGGTAAATCTTGTGCTGCTGCTG 1036
Db 755 CTGCTAAAGCTATCGGCTTCTGTTATCCAGAAATGAAATGGTAAATCTTGTGCTGCTGCTG 814
QY 1037 AACGTGTTCTGTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1096
Db 815 AACGTGTTCTGTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 874
QY 1097 TTTCTGTTGAGCAAAATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1156
Db 875 TTTCTGTTGAGCAAAATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 934
QY 1157 CTGAAGATCCAAATGTTCTTTCAGATATCGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1216
Db 935 CTGAAGATCCAAATGTTCTTTCAGATATCGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 994
QY 1217 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTTAAAGTTGTTATCATG 1276
Db 995 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTTAAAGTTGTTATCATG 1054
QY 1277 ATGCAATGAATGCTTTTACCTGCTCACTGCTCACTGCTCACTGCTCACTGCTCACTGCT 1336
Db 1055 ATGCAACGAAATGCTTTTACCTGCTCACTGCTCACTGCTCACTGCTCACTGCTCACTGCT 1114
QY 1337 TCGCTAAATAA 1347
Db 1115 TTGCTAAATAA 1125

RESULT 6
ID AAC66140
XX AAC66140 standard; DNA; 1125 BP.
AC AAC66140;

DT 13-FEB-2001 (first entry)
DE Streptococcal plasmin receptor gene plr.

KW Plasmin receptor; plr; immune response; vaccination; antibacterial;
KW Streptococcal infection; antihelminthic; ds.
XX Streptococcus pyogenes.

XX Key Location/Qualifiers
XX CDS 115..1125
XX /*tag= a
XX /product= "Plasmin receptor"

XX US6136323-A.
XX 24-OCT-2000.
XX 11-JUL-1994; 94US-00273247.
XX 29-MAR-1989; 89US-00330849.
XX 16-MAY-1990; 90US-00524411.
XX 10-AUG-1992; 92US-00928462.

XX (UNFL) UNIV FLORIDA RES FOUND INC.
XX PA Von Mering G, Broder C, Lottenberg R, Boyle MDP;
XX PI WFI; 2001-006210/01.
XX DR

DR P-PSDB; AAY85681.
XX Raising an immune response in a mammal, especially for vaccination
PT against group A streptococcal infections, comprises administering an
PT isolated streptococcal plasmin receptor protein.
XX Disclosure; Col 25-28; 17pp; English.
XX Polynucleotide sequence AAC66140 encodes the Streptococcal plasmin
CC receptor protein AAY85681. The protein is used in a method for raising an
CC immune response in a mammal. The method comprises administering the
CC plasmin receptor protein. The method is useful as a vaccination against
CC group A Streptococcal infections and potentially against a broad range of
CC infections associated with pathogens expressing glyceraldehyde-3-
CC phosphate dehydrogenase (GAPDH) or GAPDH-like proteins at their cell
CC surfaces. The vaccine has antibacterial and antihelminthic activity
XX
SQ Sequence 1125 BP; 336 A; 222 C; 234 G; 333 T; 0 U; 0 Other;

Query Match 58.0%; Score 780.6; DB 4; Length 1125;
Best Local Similarity 94.8%; Pred. No. 8.9e-198;
Matches 807; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 497 TCGATGGTACTGTTGAAGTTAAAGAGGTGGATTTCGAAGTTAAACGGTCAATTTGTTAAAG 556
Db 275 TTGATGGAACAGTTCGAAGTTAAAGAGGTGGATTTCGAAGTTAAACGGTCAATTTGTTAAAG 334
QY 557 TTTCTGCTGAAACGCAACGCAACATGATGCTGGGCTACTGATGCGGTAGAAATCGTTC 616
Db 335 TTTCTGCTGAAACGCTGATCCAGAAAACATCGAATCGGGCAACTGATGGGGTTGAAATCGTTC 394
QY 617 TTGAAGCAACTAGTTTCTTCTGCTAAAGAAAGAGTCTGCTGAAAACACATCTACATGCTAACG 676
Db 395 TTGAAGCAACTAGTTTCTTCTGCTAAAGAAAGAGTCTGCTGAAAACACATCTACATGCTAACG 454
QY 677 GTGCTAAAGAGTTGTTATCATCAGCTCTCTGTTGAAACGAGCTTAAACAGTTGTTTCA 736
Db 455 GTGCTAAAGAGTTGTTATCATCAGCTCTCTGTTGAAACGAGCTTAAACAGTTGTTTCA 514
QY 737 ACACCTAACACGACATCTTTCGAGCTGCTGAAACAGTTATCTCAGGTGCTTCATGTACTA 796
Db 515 ACACCTAACACGACATCTTTCGAGCTGCTGAAACAGTTATCTCAGGTGCTTCATGTACTA 574
QY 797 CAACCTGTTAGCTCTCTATGCTTAAAGCTCTTACGATGCTATTTGGTATCCAAAAGGTC 856
Db 575 CAACCTGTTAGCTCTCTATGCTTAAAGCTCTTACGATGCTATTTGGTATCCAAAAGGTC 634
QY 857 TTATGACTACAAATCCACGCTTATATCTGCTGACCAATGAAATGGTAAATCTTGTGCTGCTG 916
Db 635 TTATGACTACAAATCCACGCTTATATCTGCTGACCAATGAAATGGTAAATCTTGTGCTGCTG 694
QY 917 GTGGTGACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 695 GTGGTGACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 754
QY 977 CTGCTAAAGCTATCGGCTTCTGTTATCCAGAAATGAAATGGTAAATCTTGTGCTGCTGCTG 1036
Db 755 CTGCTAAAGCTATCGGCTTCTGTTATCCAGAAATGAAATGGTAAATCTTGTGCTGCTGCTG 814
QY 1037 AACGTGTTCTGTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1096
Db 815 AACGTGTTCTGTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 874
QY 1097 TTTCTGTTGAGCAAAATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1156
Db 875 TTTCTGTTGAGCAAAATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 934
QY 1157 CTGAAGATCCAAATGTTCTTTCAGATATCGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1216
Db 935 CTGAAGATCCAAATGTTCTTTCAGATATCGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 994
QY 1217 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTTAAAGTTGTTATCATG 1276

Db 995 CAACTCAAACTAAAGTAAGTAAGTGAACGATCAACATTTGGTTAAAGTTGATCATGCT 1054
QY 1277 ATGACATGAAGTCTTACACGCTCAACTGTTGTCACATTTGAGTATTTGCAAAA 1336
Db 1055 ATGACACGAAATGCTTACACTGCTCAACTTGTACGTACTCTTGAGTACTTCGCAAAA 1114
QY 1337 TCCTAAATAA 1347
Db 1115 TTGCTAAATAA 1125
RESULT 7
ABN70591
ID ABN70591 standard; DNA; 1008 BP.
XX
AC ABN70591;
XX
DT 01-JUL-2002 (first entry)
XX
XX Streptococcus polynucleotide SEQ ID NO 9095.
DE
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS Streptococcus pyogenes.
XX
XX WO200234771-A2.
XX
XX 02-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-GB004789.
XX
XX 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
XX (CHIR-) CHIRON SPA.
PA
XX (GENO-) INST GENOMIC RES.
XX
XX Telford J, Maignani V, Margarit Y Rosi, Grandi G, Fraser C;
PI Tettelin H;
PI
XX WPI: 2002-352536/38.
DR P-PSDB; ABP29960.
XX
XX New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
XX Claim 7; Page 4030; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
XX Sequence 1008 BP; 298 A; 206 C; 213 G; 291 T; 0 U; 0 Other;

Query Match 57.7%; Score 777.6; DB 6; Length 1008;

Best Local Similarity 94.8%; Pred. No. 5.4e-197;
Matches 804; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 497 TCGATGCTACTGTTGAAAGTTAAAGAGAGTGAATCGAAGTTAAACGGTCAATTTGTTAAAG 556
Db 161 TTGATGGAACAGTTTGAAGTTAAAGAGAGTGAATTTGAAGTAAACGAAACCTTCATCAAG 220
QY 557 TTTCTGCTGACGCGAACACAGCAACATTTGACTGGCTACTGATGGGTAGATAATCGTTC 616
Db 221 TTTCTGCTGACGCGAACACAGCAACATTCGACTGGCAACATCGATGGGGTTGAAATCGTTC 280
QY 617 TTGAAGCAAACTAGTTTCTTTGCTAAAAAGAGTGTCTGAAAAACACTTACATCTCTAAACG 676
Db 281 TTGAAGCAAACTAGTTTCTTTGCTAAAAAGAGTGTCTGAAAAACACTTACATCTCTAAACG 340
QY 677 GTGCTAAAAAGTTGTTATACACAGCTCTCTGGTGAAGAGAGTGTCTGAAAAAGTCTTTTCA 736
Db 341 GTGCTAAAAAGTTGTTATACACAGCTCTCTGGTGAAGAGAGTGTCTGAAAAAGTCTTTTCA 400
QY 737 ACCTAACACAGCACTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTTACTA 796
Db 401 ACCTAACACAGCACTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTTACTA 460
QY 797 CAAACTGTTTGTCTCTATGCTTAAAGCTCTTCAACGATGCAATTTGGTATCCAAAAAGGTC 856
Db 461 CAAACTGTTTGTCTCTATGCTTAAAGCTCTTCAACGATGCAATTTGGTATCCAAAAAGGTC 520
QY 857 TTATGACTCAATCCACGCTTATCTGAGTGTGACCAATGATCTTGACGGACCAACCGTG 916
Db 521 TTATGACTCAATCCACGCTTATCTGAGTGTGACCAATGATCTTGACGGACCAACCGTG 580
QY 917 GTGGTGACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 581 GTGGTGACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
QY 977 CTGCTAAGCTATCGGTCTGTTTATCCAGAAATGATGTTAACTTGTGTTAACTCTTTGATAA 1036
Db 641 CTGCTAAGCTATCGGTCTGTTTATCCAGAAATGATGTTAACTTGTGTTAACTCTTTGATAA 700
QY 1037 AACGTGTTCTGTTCCAACTGGATCACTAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1096
Db 701 AACGTGTTCTGTTCCAACTGGATCACTAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 760
QY 1097 TTTCTGTTGACAAATCAACGCTGCTGATGAAAGCTGCTTCAACAGAGTTTCGGTTACA 1156
Db 761 TTTCTGTTGACAAATCAACGCTGCTGATGAAAGCTGCTTCAACAGAGTTTCGGTTACA 820
QY 1157 CTGAAGATCAATGTTTCTTTCAGATATCGTAGGCTGTCTACACGGTTCATTTGTTGACG 1216
Db 821 CTGAAGATCAATGTTTCTTTCAGATATCGTAGGCTGTCTACACGGTTCATTTGTTGACG 880
QY 1217 CAATCAAACTAAAGTTATGGAAGTTGACGATCACAATTTGTTAAAGTTGATCATGTT 1276
Db 881 CAATCAAACTAAAGTTATGGAAGTTGACGATCACAATTTGTTAAAGTTGATCATGTT 940
QY 1277 ATGACATGAATGTTCTTACACTGCTCAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1336
Db 941 ATGACATGAATGTTCTTACACTGCTCAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1000
QY 1337 TCCTAAA 1344
Db 1001 TTGCTAAA 1008
RESULT 8
ABN69737
ID ABN69737 standard; DNA; 1035 BP.
XX
AC ABN69737;
XX
XX 01-JUL-2002 (first entry)
XX
XX Streptococcus polynucleotide SEQ ID NO 7387.

XX Novel GapC multiple epitope fusion polypeptide comprising antigenic
PT determinant of *Streptococcus dysgalactiae*, *S. agalactiae*, *S. uberis*, *S.*
PT *parvauberis*, or *S. iniae* GapC protein useful for treating mastitis in
PT vertebrates.
XX
PS Example 3; Fig 2A-B; 116pp; English.
XX
CC The present sequence is that of the coding region of the gene encoding
CC the GapC plasmid binding protein, AgalGapC (see AM50666), of
CC *Streptococcus agalactiae* ATCC 27541. The gene was obtained by PCR
CC amplification of chromosomal DNA using the primers given in AB91333-34.
CC The invention relates to novel GapC multiple epitope fusion proteins that
CC comprise epitopes from 1 or more of *Streptococcus dysgalactiae*,
CC *Streptococcus agalactiae*, *Streptococcus parvauberis*, *Streptococcus uberis*
CC and *Streptococcus iniae* (see AM50665-69). A claimed example is PolyGap4
CC (see AM50664). The multiple epitope fusion proteins are produced using
CC claimed host-vector systems and used in claimed vaccines for treating or
CC preventing a bacterial infection in a vertebrate, especially a
CC streptococcal infection, and particularly mastitis. They are also used in
CC claimed methods of detecting *Streptococcus* antibodies. The multiple
CC epitope proteins are capable of eliciting broad immunity against a
CC variety of streptococcal infections while minimising the number of
CC antigens present in the final formulation and concomitantly reducing
CC production costs
XX
SQ Sequence 1011 BP; 305 A; 204 C; 213 G; 289 T; 0 U; 0 Other;

Query Match 52.1%; Score 702.2; DB 6; Length 1011;
Best Local Similarity 89.1%; Pred. No. 7e-177;
Matches 758; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 497 TCGATGGTACTGTTGAAGTTAAAGAGGTGATTCGAAAGTTAAACGGTCAATTGTTAAAG 556
Db 161 TCGACGGTACTGTTGAAGTTAAAGAGGTGATTCGAAAGTTAAACGGTCAATTGTTAAAG 220
QY 557 TTTCTGCTGAACGGCAACCAACATTTGCTGGCTACTGATGGGTGAGAAATCGTTC 616
Db 221 TTTCTGCTGAACGGCAACCAACATTTGCTGGCTACTGATGGGTGAGAAATCGTTC 280
QY 617 TTGAAGCAACTAGTTTCTTCTTAAAGAAAGTCTGTAAGAAACACTTACATGCTAACG 676
Db 281 TTGAAGCAACTAGTTTCTTCTTAAAGAAAGTCTGTAAGAAACACTTACATGCTAACG 340
QY 677 GTGCTTAAAGATGTTTATACAGCTCTCTGTGGAACAGGTTAAACAGTGTGTTTCA 736
Db 341 GTGCTTAAAGATGTTTATACAGCTCTCTGTGGAACAGGTTAAACAGTGTGTTTCA 400
QY 737 ACATTAACCCAGCATCTTTCAGCGTACTGAAACAGTATCTCAGGTGCTTCACTACTA 796
Db 401 ACATTAACCCAGCATCTTTCAGCGTACTGAAACAGTATCTCAGGTGCTTCACTACTA 460
QY 797 CAACACTGTTAGTCTTATGCTTAAAGTCTTACAGATGCAATTTGGTATCCAAAAGTTC 856
Db 461 CAACACTGTTAGTCTTATGCTTAAAGTCTTACAGATGCAATTTGGTATCCAAAAGTTC 520
QY 857 TTATGACTTACAACTCACGCTTATCTGTGACCAAAATGATCTTGCAGGACCAACCGGTG 916
Db 521 TGATGACTTATCAACGATACACTGTGTGACCAAAATGATCTTGCAGGACCAACCGGTG 580
QY 917 GTGTGACCTTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 581 GTGTGACCTTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
QY 977 CTGTGAAGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1036
Db 641 CTGTGAAGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 700
QY 1037 AACGTGTTCTGTTCCAACTGGATCAGTAACCTAGTGTGGTTGTAACCTCTTGATAAAACG 1096
Db 701 AACGTGTTCTGTTCCAACTGGATCAGTAACCTAGTGTGGTTGTAACCTCTTGATAAAACG 760
QY 1097 TTTCTGTTGACGAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTTCGGTTACA 1156

Db 761 TAACTGCTGAAGAGTAATATGAGCTATGAAGCAGCAGCTAACGATTTCATACGGTTATA 820
QY 1157 CTGAAGATCAATGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1216
Db 821 CTGAAGATCAATGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 880
QY 1217 CAACCTCAAACTAAAGTTATGGAAGTTGAGGATCAACAATGGTTAAAGTTGTATCATGGT 1276
Db 881 CTAACTCAAACTAAAGTTCAAACTGTTGAGGTAACCAATGGTTAAAGTTGTATCATGGT 940
QY 1277 ATGACAAATGAATGTTCTTACACTGCTCAACTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1336
Db 941 ACGATAACGAAATGTCATACACTTCAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1000
QY 1337 TCGCTAAATAA 1347
Db 1001 TCGCTAAATAA 1011

RESULT 10
ABA91249
ID ABA91249 standard; DNA; 1011 BP.
XX
AC ABA91249;
XX
DT 04-APR-2002 (first entry)
XX
DE *Streptococcus agalactiae* gapC gene.
XX
KW GapC; plasmin-binding protein; infection; mastitis; vaccine; diagnosis;
KW gene therapy; gene; ds.
XX
OS *Streptococcus agalactiae*.
XX
PN WO2001196381-A2.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-CA000838.
XX
PR 12-JUN-2000; 2000US-0211022P.
XX
PA (UUSA-) UNIV SASKATCHEWAN.
XX
PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
XX
DR WPI; 2002-130725/17.
XX
DR P-PSDB; AAM50640.
XX
PT Novel isolated GapC protein of *Streptococcus dysgalactiae*, *S. agalactiae*,
PT *S. uberis*, *S. parvauberis*, or *S. iniae*, useful as vaccine component for
PT treating streptococcal infection which causes mastitis in vertebrates.
XX
PS Example 1; Fig 2A-B; 107pp; English.
XX
CC The present sequence is that of the coding region of the GapC plasmid
CC binding protein gene of *Streptococcus agalactiae* ATCC 27541. The gene was
CC amplified from chromosomal DNA by PCR using the primers given in AB91253
CC -54. The PCR product was cloned into expression vector pE15B, creating
CC plasmid pMF52ic. The gene encodes a GapC plasmid binding protein (see
CC AAM50640) that is capable of eliciting an immune response in a
CC vertebrate. The invention provides the GapC genes and proteins of 5
CC *Streptococcus* species, as well as recombinant vectors, host cells and
CC vaccine compositions comprising GapC polynucleotides or proteins. The
CC vaccines are used to treat or prevent a bacterial infection, especially a
CC streptococcal infection, and mastitis in particular (claimed). A
CC polynucleotide encoding a GapC protein is used in a claimed method of
CC treating or preventing a bacterial infection, such as a streptococcal
CC infection, especially mastitis. *S. agalactiae* is a common pathogen
CC associated with mastitis in cattle, horse, sheep and goat. It also causes
CC septicemia, meningitis, bacteraemia, impetigo, arthritis, urinary tract
CC infections, abscesses, spontaneous abortion, etc


```
XX SQ Sequence 1011 BP; 305 A; 204 C; 213 G; 289 T; 0 U; 0 Other;
Query Match 52.1%; Score 702.2; DB 6; Length 1011;
Best Local Similarity 89.1%; Pred. No. 7e-177;
Matches 758; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 497 TCGATGCTACTGTTGAAGTTAAAGAGGTGATTCGAAGTTAAGCTCAATTTGTTAAAG 556
DB 161 TCGACGGTACTGTTGAAGTTAAAGAGGTGATTCGAAGTTAAGCTCAATTTGTTAAAG 220
QY 557 TTTCTGCTGAACCGAACCAGCAACATTCAGCTGGCTACTGATGGCTAGAAATCGTTTC 616
DB 221 TTTCTGCTGAACCGAACCAGCAACATTCAGCTGGCTACTGATGGCTAGAAATCGTTTC 280
QY 617 TTGAAGCAACTAGTTTCTTTGCTTAAAGAGAGCTGCTGAAAGCACTTACATGCTAAAG 676
DB 281 TTGAAGCAACTAGTTTCTTTGCTTAAAGAGAGCTGCTGAAAGCACTTACATGCTAAAG 340
QY 677 GTGCTAAAAAGTTGTTTATCACAGCTCTCTGCTGAAAGCACTTAAAGCACTTAAAG 736
DB 341 GTGCTAAAAAGTTGTTTATCACAGCTCTCTGCTGAAAGCACTTAAAGCACTTAAAG 400
QY 737 AACTAACACGCAATTCCTTGAAGGTACTGAAAGCACTTACAGGTGCTTCAATGCTACTA 796
DB 401 AACTAACACGCAATTCCTTGAAGGTACTGAAAGCACTTACAGGTGCTTCAATGCTACTA 460
QY 797 CAAGCTGTTAGCTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 856
DB 461 CAAGCTGTTAGCTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 520
QY 857 TTATGACTACAATCCAGCTTATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 916
DB 521 TGATGACTACTATCCAGCTATACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580
QY 917 GTGCTGACCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
DB 581 GTGCTGACCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
QY 977 CTGCTAAAGCTATCGGCTTCTGTTATCCAGAAATGAAATGTAATGTAATGTAATGTA 1036
DB 641 CTGCTAAAGCTATCGGCTTCTGTTATCCAGAAATGTAATGTAATGTAATGTAATGTA 700
QY 1037 AACGTGTTCTGTTCCAACTGGAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1096
DB 701 AACGTGTTCTGTTCCAACTGGAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 760
QY 1097 TTTCTGTTGACGAAATCAAGCTGCTATGAAAGCTGCTTCAAGCAAGCTTTCGCTTACA 1156
DB 761 TAACTGTCGAAGAGTAATGCAAGCTATGAAAGCAAGCTTTCGCTTACA 820
QY 1157 CTGAAGTCCAAATGTTTCTTCAATATCGTAGCGTGTGATACGTTTCAATGTTGAGC 1216
DB 821 CTGAAGTCCAAATGTTTCTTCAATATCGTAGCGTGTGATACGTTTCAATGTTGAGC 880
QY 1217 CAATCTAACTAAAGTTATGGAAGTTCAGGATGCAATGTTTAAAGTTGTAATGTTGTA 1276
DB 881 CTACTCAAACTAAAGTTCAAACTGTTGAGGTAACCAATGTTTAAAGTTGTTTCAATG 940
QY 1277 ATGCAATGAAGTGTCTTACACTGCTCAACTGTTGTTGTAACACTGTTGTTGTAACAA 1336
DB 941 ACATACGAAATGTCTACACTTCAAACTGTTGTTGTAACACTGTTGTTGTAACAA 1000
QY 1337 TCGCTAAATTA 1347
DB 1001 TCGCTAAATTA 1011

RESULT 11
ABN71527_17/c
Continuation (18 of 22) of ABN71527 from base 1700001 (Streptococcus polynucleotide SEQ
WP Sequence split into 22 fragments LOCUS ABN71527 Accession ABN71527
WP Fragment Name Begin End
```

```
WP ABN71527_00 1 110000
WP ABN71527_01 100001 210000
WP ABN71527_02 200001 310000
WP ABN71527_03 300001 410000
WP ABN71527_04 400001 510000
WP ABN71527_05 500001 610000
WP ABN71527_06 600001 710000
WP ABN71527_07 700001 810000
WP ABN71527_08 800001 910000
WP ABN71527_09 900001 1010000
WP ABN71527_10 1000001 1110000
WP ABN71527_11 1100001 1210000
WP ABN71527_12 1200001 1310000
WP ABN71527_13 1300001 1410000
WP ABN71527_14 1400001 1510000
WP ABN71527_15 1500001 1610000
WP ABN71527_16 1600001 1710000
WP ABN71527_17 1700001 1810000
WP ABN71527_18 1800001 1910000
WP ABN71527_19 1900001 2010000
WP ABN71527_20 2000001 2110000
WP ABN71527_21 2100001 2155561

Query Match 51.9%; Score 699; DB 6; Length 110000;
Best Local Similarity 88.8%; Pred. No. 4.1e-175;
Matches 756; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 497 TCGATGCTACTGTTGAAGTTAAAGAGGTGATTCGAAGTTAAGCTCAATTTGTTAAAG 556
DB 30602 TCGACGGTACTGTTGAAGTTAAAGAGGTGATTCGAAGTTAAGCTCAATTTGTTAAAG 30543
QY 557 TTTCTGCTGAACCGAACCAGCAACATTCAGCTGGCTACTGATGGCTAGAAATCGTTTC 616
DB 30542 TTTCTGCTGAACCGAACCAGCAACATTCAGCTGGCTACTGATGGCTAGAAATCGTTTC 30483
QY 617 TTGAAGCAACTAGTTTCTTTGCTTAAAGAGAGCTGCTGAAAGCACTTACATGCTAAAG 676
DB 30482 TTGAAGCAACTAGTTTCTTTGCTTAAAGAGAGCTGCTGAAAGCACTTACATGCTAAAG 30423
QY 677 GTGCTAAAAAGTTGTTTATCACAGCTCTCTGCTGAAAGCAAGCTTAAAGCACTTAAAG 736
DB 30422 GTGCTAAAAAGTTGTTTATCACAGCTCTCTGCTGAAAGCAAGCTTAAAGCACTTAAAG 30363
QY 737 ACATTAACACGCACTTCTTGAAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
DB 30362 ACATTAACACGCACTTCTTGAAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 30303
QY 797 CAAGCTGTTAGCTTCTTATGCTTCAAGCTCTTCAAGCTGCTGCTGCTGCTGCTGCTGCT 856
DB 30302 CAAGCTGTTAGCTTCTTATGCTTCAAGCTCTTCAAGCTGCTGCTGCTGCTGCTGCTGCT 30243
QY 857 TTATGACTACAATCCAGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 916
DB 30242 TTATGACTACAATCCAGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 30183
QY 917 GTGCTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
DB 30182 GTGCTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 30123
QY 977 CTGCTAAAGCTATCGGCTTCTTATCCAGAAATGAAATGTAATGTAATGTAATGTAATG 1036
DB 30122 CTGCTAAAGCTATCGGCTTCTTATCCAGAAATGAAATGTAATGTAATGTAATGTAATG 30063
QY 1037 AACGTGTTCTGTTTCCAACTGGAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1096
DB 30062 AACGTGTTCTGTTTCCAACTGGAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 30003
QY 1097 TTTCTGTTGACGAAATCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156
DB 30002 TAACTGTCGAAGAGTAATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 29943
QY 1157 CTGAAGATCCAAATGTTTCTTCAAGATATCGTAGGCTGCTGCTGCTGCTGCTGCTGCT 1216
```


Db 29942 CTGAAGATCCATCGTATCATCTGATATCGTTGGTATTTCATACCGTTCAATGTTTGATG 29883
 Qy 1217 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCAAAATGGTTAAAGTTGTATCATGGT 1276
 Db 29882 CTACTCAAACTAAAGTTTCAAACTGTTGACGGTAAACCAATGGTTAAAGTTGTTTCATGGT 29823
 Qy 1277 ATGACAATGAATGCTTTACACTGCTCAACTGTTGCGTACACTTGAGTATTTCGCAAAA 1336
 Db 29822 ACGATAACGAAATGTCATACACTTCACAACCTGTTTCGTACACTTGAATACTTTGCAAAA 29763
 Qy 1337 TCGCTAAATAA 1347
 Db 29762 TTGCAAAATAA 29752

RESULT 12

ABN71389
 ID ABN71389 standard; DNA; 1008 BP.

XX AC ABN71389;
 XX XX
 DT 01-JUL-2002 (first entry)
 XX Streptococcus polynucleotide SEQ ID NO 10691.
 DE Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
 KW antinflammatory; infection; vaccine; meningitis; gene therapy; ds.
 XX OS Streptococcus agalactiae.

XX WO200234771-A2.
 PN 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB004789.
 XX 27-OCT-2000; 2000GB-00026333.
 PR 24-NOV-2000; 2000GB-00028727.
 PR 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
 PI Tettelin H;

XX WPI; 2002-352536/38.
 DR P-PSDB; ABP30758.

XX New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.

XX Claim 7; Page 4178-4179; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins

XX SQ Sequence 1008 BP; 306 A; 203 C; 211 G; 288 T; 0 U; 0 Other;
 Query Match 51.7%; Score 696; DB 6; Length 1008;
 Best Local Similarity 88.8%; Pred No. 3.2e-175;
 Matches 753; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
 Qy 497 TCGATGGTACTGTGAGTTTAAAGAGGTGGATTCGAAAGTTAAACGGTCAATTTGTTAAAG 556
 Db 161 TCGACGGTACTGTGAGTTTAAAGAGGTGGATTCGAAAGTTAAACGGTCAATTTGTTAAAG 220
 Qy 557 TTTCTGCTGAACCGCAACGAGCAAAACATTGACTGGGCTACTGATGGCTAGAGTAATCGTTC 616
 Db 221 TTTCTGCTGAACCGCAACGAGCAAAACATTGACTGGGCTACTGATGGCTAGAGTAATCGTTC 280
 Qy 617 TTGAAGCAACTAGTCTTTCTTTGCTTAAAAAGAAAGCTCTGAAAAACACTTACATGCTAAAG 676
 Db 281 TTGAAGCAACTAGTCTTTCTTTGCTTAAAAAGAAAGCTCTGAAAAACACTTACATGCTAAAG 340
 Qy 677 GTGCTAAAAAGTTGTTATCACAGCTCTGCTGGTGGAAACGACGTTAAAAACAGTTGTTTCA 736
 Db 341 GTGCTAAAAAGTTGTTATCACAGCTCTGCTGGTGGAAACGACGTTAAAAACAGTTGTTTCA 400
 Qy 737 ACATAACACGACATTTCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGACTA 796
 Db 401 ACATAACACGACATTTCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGACTA 460
 Qy 797 CAACCTGTTAGCTCTCTATGGCTTAAAGCTCTTCCAGCATGCTTTGGTATCCAAAAAGTTC 856
 Db 461 CAACCTGTTAGCTCTCTATGGCTTAAAGCTCTTCCAGCATGCTTTGGTATCCAAAAAGTTC 520
 Qy 857 TTATGACTACAATCCACGCTTTATCTGCTGGTGGACCAATGATCTTTGACGGACCAACCGTG 916
 Db 521 TGATGACTACTATCCACGCTTACCTGCTGGTGGACCAATGATCTTTGACGGACCAACCGTG 580
 Qy 917 GTGGTGACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
 Db 581 GTGGTGACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
 Qy 977 CTGCTAAAGCTATCGGCTCTGTTTATCCAGAAATGATGTTAACTTGGTGGTGGTGCAC 1036
 Db 641 CTGCTAAAGCTATCGGCTCTGTTTATCCAGAAATGATGTTAACTTGGTGGTGGTGCAC 700
 Qy 1037 AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTGGTGTGTAATCTTTGATAAAAACG 1096
 Db 701 AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTGGTGTGTAATCTTTGATAAAAACG 760
 Qy 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTTTCGGTTACA 1156
 Db 761 TAACTGTGGAAGAAAGTAAATGCGAGCTATGAAAGCAGCAGCTAACGATTATACGGTTATA 820
 Qy 1157 CTGAAGATCCAAATGTTTCTTTCAGATATCGTAGGCGTGTGATACGGTTTCATTTTGAAG 1216
 Db 821 CTGAAGATCCAAATGTTTCTTTCAGATATCGTAGGCGTGTGATACGGTTTCATTTTGAAG 880
 Qy 1217 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCACAANTGGTTAAAGTTGATCATGGT 1276
 Db 881 CTACTCAAACTAAAGTTCAAACTGTTGACGGTTAAACCAATGGTTAAAGTTGTTTCATGGT 940
 Qy 1277 ATGCAATGAAATGTTTACACCTGCTCAACTGCTTCTGCTACACTTTGAGTATTTTCAAAA 1336
 Db 941 ACGATACGAAATGTTATACACTTTCACACTTGTTCGCTACACTTGTGATACACTTTGCAAAA 1000
 Qy 1337 TCGCTAAA 1344
 Db 1001 TTGCAAAA 1008

RESULT 13

AAS55858
 ID AAS55858 standard; DNA; 1080 BP.
 XX
 AC AAS55858;

XX DT 13-FEB-2002 (first entry)
 XX DE Streptococcus pneumoniae DNA for cellular proliferation protein #429.
 XX KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
 XX KW antibacterial; drug design.
 XX OS Streptococcus pneumoniae.
 XX FN WO200170955-A2.
 XX PD 27-SEP-2001.
 XX PF 21-MAR-2001; 2001WO-US009180.
 XX PR 21-MAR-2000; 2000US-0191078P.
 XX PR 23-MAY-2000; 2000US-0206848P.
 XX PR 26-MAY-2000; 2000US-0207727P.
 XX PR 23-OCT-2000; 2000US-0242578P.
 XX PR 27-NOV-2000; 2000US-0253625P.
 XX PR 22-DEC-2000; 2000US-0257931P.
 XX PR 16-FEB-2001; 2001US-0269308P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 XX PI Yamamoto RT, Xu HH;
 XX PR WPI; 2001-611495/70.
 XX PR P-PSDB; AAU37999.
 XX PT New polynucleotides for the identification and development of
 XX PT antibiotics, comprise sequences of antisense nucleic acids.
 XX PS Claim 27; SEQ ID NO 9495; 51pp; English.
 XX CC The invention relates to antisense inhibitors of genes essential to
 XX CC prokaryotic cellular proliferation, their use in identifying the genes,
 XX CC their use in the discovery of novel antibiotics, the essential genes
 XX CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*
 XX CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,
 XX CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also
 XX CC useful for the identification of potential new targets for antibiotic
 XX CC development. The antisense nucleic acids can also be used to identify
 XX CC proteins used in proliferation, to express these proteins, and to obtain
 XX CC antibodies capable of binding to the expressed proteins. The proteins can
 XX CC be used to screen compounds in rational drug discovery programmes. The
 XX CC antisense nucleic acid sequence is also useful to screen for homologous
 XX CC nucleic acids which are required for cell proliferation in a wide variety
 XX CC of organisms. The present sequence encodes an essential prokaryotic
 XX CC cellular proliferation protein. Note: The sequence data for this patent
 XX CC did not form part of the printed specification, but was obtained in
 XX CC electronic format directly from WIPO at
 XX CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 1080 BP; 320 A; 214 C; 234 G; 312 T; 0 U; 0 Other;
 Query Match 50.7%; Score 683; DB 4; Length 1080;
 Best Local Similarity 88.5%; Pred. No. 9.5e-172;
 Matches 753; Conservative 0; Mismatches 95; Indels 3; Gaps 1;
 QY 497 TCGATGTTACTGTTGAAGTTAAAGAGGTGATTCGAAGTTACGGTCAATTTGTTAAAG 556
 DB 233 TCGACGGTACTGTTGAAGTTAAAGAGGTGATTCGAAGTTACGGTCAATTTGTTAAAG 292
 QY 557 TTTCGTGTGAACCGAACCGAACAAATTCAGTGGGCTACTGATGGCGTAGAAATCGTTC 616
 DB 293 TTTCGTGTGAACCGAACCGAACAAATTCAGTGGGCTACTGATGGGCTAGAAATCGTTC 352
 QY 617 TTGAAGCACTAGTTCTTTGCTTAABAAAGAGCTGCTGAACCACTTACATGCTAACG 676
 DB 353 TTGAAGCACTAGTTCTTTGCTTAABAAAGAGCTGCTGAACCACTTACATGCTAACG 409

QY 677 GTGCTAAAAAGTTGTTATACAGCTCCTGGTGGAAACGAGCTTAAAAAGTTGTTTCA 736
 DB 410 GAGCTAAAAAGTTGTTATACAGCTCCTGGTGGAAACGAGCTTAAAAAGTTGTTTCA 469
 QY 737 ACACCTAACACGACATCTTACGCTACTGAAACAGTTATCTCAGTGCCTTCATGACTA 796
 DB 470 ACACCTAACACGACATCTTACGCTACTGAAACAGTTATCTCAGTGCCTTCATGACTA 529
 QY 797 CAAACTGTTTAGCTCCTATGCTTAAAGCTCTTACAGATGCTATTTGGTATCCAAAAAGGTC 856
 DB 530 CAAACTGCTTGGCTCCAATGGCTAAAGCTCTTCAAGATAACTTTTGGTGTGTTGAAGAT 589
 QY 857 TTATGACTACAAATCCACGCTTATCTGCTGACCAAAATGATCTTGCACGACACACCGTG 916
 DB 590 TGATGACTACTATCCACGCTTACACTGCTGACCAAAATGATCTTGCACGACACACCGTG 649
 QY 917 GTGGTGACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 976
 DB 650 GTGGTGACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 709
 QY 977 CTGCTAAAGCTATCGTCTTGTATCCAGAAATGGAATGGTAAAGTTGATGGTGTGCAC 1036
 DB 710 CTGCAAAAGCTATCGTCTTGTATCCAGAAATGGAATGGTAAAGTTGATGGATCTGCAC 769
 QY 1037 AACGTGTTCTCTTCCAACTGGATCAGTAACTGAGTTGGTTGTTAACTCTTGATAAAACG 1096
 DB 770 AACGCTTCCAACTCCAACTGGATCAGTAACTGAGTTGGTTGTTAACTCTTGATAAAACG 829
 QY 1097 TTTCTGTTGACGAAATCAACGCTGCTGATGAAAGCTGTTCAAAACGACAGTTTCGTTTACA 1156
 DB 830 TTACTGTTGATGAAGTGAACGACGCTATGAAAGCAGCTTCAACGGAATCATACGTTTACA 889
 QY 1157 CTGAAGATCCAAATTTCTTCCAGATATCGTAGGCTGTCATACGTTTCATTTGTTGACG 1216
 DB 890 CAGAAAGTCCAAATCTGATCTTCCAGATATCGTAGGCTGTCATACGTTTCATTTGTTGACG 949
 QY 1217 CAACTCAAACTAAAGTTATGGAAGTTGACGATCACAATTCGTTAAAGTTGATCATCGT 1276
 DB 950 CAACTCAAACTAAAGTTCTTGACGTTGACGCTGAAACAAATTCGTTAAAGTTGATCATCGT 1009
 QY 1277 ATGCAAGTGAATGCTTCTTACACTGCTCAACTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1336
 DB 1010 AGCAACGAAATGTCATACACTGCAAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1069
 QY 1337 TCGCTAAATAA 1347
 DB 1070 TTGCTAAATAA 1080
 RESULT 14
 ID ABA91332
 XX ABA91332 standard; DNA; 1011 BP.
 XX AC ABA91332;
 XX DT 08-APR-2002 (first entry)
 XX DE Streptococcus iniae gapC plasmin binding protein gene.
 XX KW InnaseGapC; GapC; plasmin binding protein; epitope; infection; vaccine;
 XX KW immunisation; mastitis; therapy; gene; ds.
 XX OS Streptococcus iniae.
 XX FN WO200196379-A2.
 XX PD 20-DEC-2001.
 XX PF 11-JUN-2001; 2001WO-CA000836.
 XX PR 12-JUN-2000; 2000US-0211247P.
 XX

(UUSA-) UNIV SASKATCHEWAN.
 Potter AA, Perez-Casal J, Fontaine M;
 WPI; 2002-098051/13.
 P-PSDB; ABA50669.

Novel GapC multiple epitope fusion polypeptide comprising antigenic determinant of *Streptococcus dysgalactiae*, *S. agalactiae*, *S. uberis*, *S. parvauberis*, or *S. iniae* GapC protein useful for treating mastitis in vertebrates.

Example 3; Fig 5A-B; 116pp; English.

The present sequence is that of the coding region of the gene encoding the GapC plasmid binding protein, IniaeGapC (see AAM50669), of *Streptococcus iniae* strain 9117. The gene was obtained by PCR amplification of chromosomal DNA using the primers given in ABA91333-34. The invention relates to novel GapC multiple epitope fusion proteins that comprise epitopes from 1 or more of *Streptococcus dysgalactiae*, *Streptococcus agalactiae*, *Streptococcus parvauberis*, *Streptococcus uberis* and *Streptococcus iniae* (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664). The multiple epitope fusion proteins are produced using claimed host-vector systems and used in claimed vaccines for treating or preventing a bacterial infection in a vertebrate, especially a streptococcal infection, and particularly mastitis. They are also used in claimed methods of detecting *Streptococcus* antibodies. The multiple epitope proteins are capable of eliciting broad immunity against a variety of streptococcal infections while minimising the number of antigens present in the final formulation and concomitantly reducing production costs

Sequence 1011 BP; 305 A; 191 C; 213 G; 302 T; 0 U; 0 Other;

Query Match 50.3%; Score 678.2; DB 6; Length 1011;
 Best Local Similarity 87.3%; Pred. No. 1.8e-170;
 Matches 743; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

497 TCGATCGTACTGTTGAAGTTAAAGAGGTGAGTTCGAAGTTAAGCGTCAATTTGTTAAAG 556
 DB 161 TTGACGGTACAGTTGAAGTTAAAGATGTTGAGTTTCGAAGTTAAGCGGAGCTTTGTTAAAG 220

557 TTCTCTGCTGAACGCAAGCAAGCAAGCAATTCAGTGGCTACTGATGCGTGAAGAAATCGTTC 616
 DB 221 TTCTGCAAGACGCAAGCAAGCAAGCAATTCAGTGGCTACTGATGCGTGAAGAAATCGTTC 280

617 TTGAAGCAACTAGTTTCTTTGCTTAAAGAAAGAGTCTCTGAAAGAAACACTTACATGCTAACG 676
 DB 281 TTGAAGCAACTAGTTTCTTTGCTTAAAGAAAGAGTCTCTGAAAGAAACACTTACATGCTAACG 340

677 GTGCTAAGAAAGTGTATATACAGCTCTGCTGGAAGCAAGCTTAAACAGTTGTTTTC 736
 DB 341 GTGCGAAAAGTGTATATACAGCTCTGCTGGAAGCAAGCTTAAACAGTTGTTTTC 400

737 ACACTAACCCAGCAATCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGCTACTA 796
 DB 401 ACACTAACCCAGCAATCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGCTACTA 460

797 CAACTGTTAGTCTCTATGCTTAAAGCTCTTCACAGTGCATTTGATGCTCAAAAGGTC 856
 DB 461 CAACTGTTAGTCTCTATGCTTAAAGCTCTTCACAGTGCATTTGATGCTCAAAAGGTC 520

857 TTATGACTACAATCCAGCTTATCTGCTGAGCAAAATGATCTTCAGCGACCAACCGGTG 916
 DB 521 TAATGACTACTATCATGTTTACATCTGTTGACCAAAATGTTCTTCAGCGACCAACCGGTG 580

917 GTGGTACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
 DB 581 GTGGTACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640

977 CTGCTAAAGCTATCGTCTCTTTTATCCAGAAATTCGAAGTAAATTCGATGCTGCTGCAC 1036
 DB 641 CTGCTAAAGCTATCGTCTCTTTTATCCAGAAATTCGAAGTAAATTCGATGCTGCTGCAC 700

1037 AACGTTTCCTGTTCCAACTCGATCAGTAAGTGGTTGGTTAACTTCTTCACTTCTTCAAAAAACG 1096
 DB 701 AACGTTTCCTGTTCCAACTCGATCAGTAAGTGGTTGGTTAACTTCTTCACTTCTTCAAAAAACG 760

1097 TTTCTGTTGACGAATCAACGCTGCTATGAAGCTGTTTCAACGACAGTTCGTTTCA 1156
 DB 761 CTTTCAAGTAAAGAAATCAATGCACTATGAAGCAGCAGCTTAACGATTCATACGTTTCA 820

1157 CTGAAGATCCAAATGTTTCTTTCAGATATCGPAGCGGTGTATACGGTTCATGTTTCAAG 1216
 DB 821 CTGAAGATGCTATGCTATCATCATGATATCGTAGGTATTTCTTACGGTTTCAATTTTATG 880

1217 CAACCTCAAACTAAAGTTATGGAAGTTGACGATCACAATTCGTTTAAAGTTGTTATCATGGT 1276
 DB 881 CTACTCAAACTAAAGTACAACTGTTGATGAATCAATTCGTTTAAAGTTGTTTCAATGGT 940

1277 ATGCAATGAATGTTTCTTACACTGCTCAACTGTTTCTGATACACTTTCGAGTATTTTCAAAAA 1336
 DB 941 ATGCAATGAATGTTTCTTACACTGCTCAACTGTTTCTGATACACTTTCGAGTATTTTCAAAAA 1000

1337 TCGCTAAATAA 1347
 DB 1001 TCGCTAAATAA 1011

RESULT 15
 ABA91252
 ID ABA91252 standard; DNA; 1011 BP.
 AC ABA91252;
 DT 04-APR-2002 (first entry)
 DE Streptococcus iniae gapC gene.
 KW GapC; plasmin-binding protein; infection; mastitis; vaccine; diagnosis;
 KW Gene therapy; gene; ds.
 OS Streptococcus iniae.
 PN WO200196381-A2.
 PD 20-DEC-2001.
 PF 11-JUN-2001; 2001WO-CA000838.
 PR 12-JUN-2000; 2000US-0211022P.
 PA (UUSA-) UNIV SASKATCHEWAN.
 PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
 DR WPI; 2002-130725/17.
 DR P-PSDB; AAM50643.
 PT Novel isolated GapC protein of *Streptococcus dysgalactiae*, *S. agalactiae*,
 PT *S. uberis*, *S. parvauberis*, or *S. iniae*, useful as vaccine component for
 PT creating streptococcal infection which causes mastitis in vertebrates.
 PS Example 1; Fig 5A-B; 107pp; English.
 CC The present sequence is that of the coding region of the GapC plasmid
 CC binding protein gene of *Streptococcus iniae* strain 9117. The gene was
 CC amplified from chromosomal DNA by PCR using the primers given in ABA91253
 CC and ABA91254. The PCR product was cloned into expression vector pE115B,
 CC creating plasmid pMF521e. The gene encodes a GapC plasmid binding protein
 CC (see AAM50643) that is capable of eliciting an immune response in a
 CC vertebrate. The invention provides the GapC genes and proteins of 5
 CC *Streptococcus* species, as well as recombinant vectors, host cells and
 CC vaccine compositions comprising GapC polynucleotides or proteins. The
 CC vaccines are used to treat or prevent a bacterial infection, especially a
 CC streptococcal infection, and mastitis in particular (Claimed). A

CC polynucleotide encoding a GapC protein is used in a claimed method of
treating or preventing a bacterial infection, such as a streptococcal
infection, especially mastitis

XX
SQ Sequence 1011 BP; 305 A; 191 C; 213 G; 302 T; 0 U; 0 Other;
Query Match 50.3%; Score 678.2; DB 6; Length 1011;
Best Local Similarity 87.3%; Pred. No. 1.8e-170;
Matches 743; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 497 TCGATGGTACTTGAAGTTAAAGAGGTGATTCGAAGTTAAACGGTCAATTGTTAAAG 556
Db 161 TTGACGGTACAGTTGAAGTTAAAGATGGTGAATCGAAGTTAAACGGAGCTTTGTTAAAG 220
QY 557 TTTCTGCTGACGCGCAACAGCAACATGCTACTGGCTACTGATGGCTAGAAATCGTTC 616
Db 221 TTTCTGCGAGCGCAACAGCAACATGCTACTGGCTACTGATGGTGTAGACATCGTTC 280
QY 617 TTGAAGCAACTAGTTTCTTGTCTAAAAAAGAGCTGTGAAAAACACCTTACATGCTAACG 676
Db 281 TTGAAGCAACAGGTTTCTTGTCTTAAAGCAGCTGTGAACACACATTCACGCTAACG 340
QY 677 GTGCTAAAAAGTTGTTATACAGCTCTCTGTGGAACGAGTTAAACAGTTGTTTCA 736
Db 341 GTGCCAAAAAGTTGTTATACAGCTCTCTGTGGAACGAGTTAAACAGTTGTTTACA 400
QY 737 ACACCTAACCAACACACATTTCTGACGGTACTGAAACAGATTATCTCAGGTGCTTCATGACTA 796
Db 401 ACACCTAACCAACAGATTCTTGTATGAACTGAAACAGTTATCTCAGGTGCTTCATGACTA 460
QY 797 CAACCTGTTTACTGCTTAAAGCTTTTCAAGATGCAATTTGGTATCCAAAAAGGTC 856
Db 461 CAACCTGTTTACTGCTTAAAGCTTAAAGCATTACAAAGATAACTTTGGTGTAAACAGGTT 520
QY 857 TTATGACTACATCCACGCTTATCTGCTGACCAAAATGATCCTTGACGGCCACACCGTG 916
Db 521 TAATGACTACTATCCATGCTTACCTGCTGACCAAAATGCTTTCGCGCCACACCGTG 580
QY 917 GTGCTGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 581 GTGCTGACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
QY 977 CTGCTAAAGCTATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1036
Db 641 CTGCTAAAGCTATCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 700
QY 1037 AACGTGTTCTGCTTCCAACTGATCAGTAACTGAGTTGGTTGTTAACTCTTGAATAAACG 1096
Db 701 AACGTGTTCTGCTTCCAACTGATCAGTAACTGAGTTGGTTGTTAACTCTTGAATAAACG 760
QY 1097 TTTCTGTTGAGAAATCAACGCTGCTGATGAAGCTGCTTCAACGACAGTTTGGTTACA 1156
Db 761 CTTCAAGTAAAGAAATCAACGCTGCTGATGAAGCAGCAGCTAACGATTCAACGTTTACA 820
QY 1157 CTGAAGATCCAAATGTTTCTTCAATATCGTAGGCTGCTATACGTTTCAATGTTTGACG 1216
Db 821 CTGAAGATGCTATGCTATCATCAGATATCGTAGGTTTCTTACGTTTCAATGTTTGATG 880
QY 1217 CAATCAAACTAAAGTTAAGAGTTGACGATCAATGGTTAAAGTTGATCATGCT 1276
Db 881 CTACTCAAACTAAAGTTAAGAGTTGATGAAATCAATGGTTAAAGTTGTTTCAATGCT 940
QY 1277 ATGCAATGAATGCTTTCATCTGCTCACTGCTCACTGCTCACTGCTCACTGCTCACTG 1336
Db 941 ATGCAATGAATGCTTTCATCTGCTCACTGCTCACTGCTCACTGCTCACTGCTCACTG 1000
QY 1337 TCGCTAAATAA 1347
Db 1001 TCGCTAAATAA 1011

Search completed: March 30, 2004, 23:30:05
Job time : 574 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 23:30:10 ; Search time 128 seconds
(without alignments)
5839.991 Million cell updates/sec

Title: US-10-650-369-21

Perfect score: 1347

Sequence: 1 atgaaaaataacacaggat.....ttgcaaaaatcgctaaataa 1347

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A COMB seq: *
2: /cgn2_6/ptodata/2/ina/5B COMB seq: *
3: /cgn2_6/ptodata/2/ina/6A COMB seq: *
4: /cgn2_6/ptodata/2/ina/6B COMB seq: *
5: /cgn2_6/ptodata/2/ina/PTCUS COMB seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1347	100.0	1347	4	US-09-878-766A-21
2	815.8	60.6	1011	4	US-09-878-766A-11
3	780.6	58.0	1125	1	US-07-928-462-1
4	780.6	58.0	1125	3	US-08-273-247-1
5	702.2	52.1	1011	4	US-09-878-766A-13
6	678.2	50.3	1011	4	US-09-878-766A-19
7	678.2	50.3	7571	4	US-08-961-527-2
8	676.6	50.2	1011	4	US-09-878-766A-15
9	660.2	49.0	1000	3	US-08-961-083-53
10	660.2	49.0	1000	4	US-09-536-784-53
11	639.8	47.5	1011	4	US-09-878-766A-17
12	495.6	36.8	1002	4	US-09-107-532A-1115
13	478	35.5	1041	4	US-09-134-000C-995
14	436.8	31.7	1047	4	US-09-134-001C-2676
15	409.2	30.4	6048	4	US-08-956-171E-128
16	277	20.6	1074	4	US-09-134-000C-824
17	269	20.0	849	4	US-09-107-532A-403
18	268.2	19.9	580073	4	US-08-545-528D-1
19	239.4	17.8	1362	6	5290690-8
20	235.6	17.5	1717	4	US-09-634-238-18
21	221.8	16.5	1928	4	US-09-674-826B-5
22	210	15.6	1023	4	US-09-543-681A-3911
23	205	15.2	1050	4	US-09-489-039A-6768
24	198.6	14.7	1008	4	US-09-489-039A-508
25	177.2	13.2	1230025	4	US-09-198-452A-1
26	169.8	12.6	35881	4	US-08-311-731A-127
27	169.2	12.6	1026	4	US-09-134-001C-946

28 160.4 11.9 4409 4 US-09-653-403-2 Sequence 2, Appli
29 160.4 11.9 4409 4 US-10-013-784-2 Sequence 2, Appli
30 160.4 11.9 4409 4 US-09-786-606-1 Sequence 1, Appli
31 160.4 11.9 4431 3 US-09-532-803-8 Sequence 8, Appli
32 159.2 11.8 4403765 3 US-09-103-840A-2 Sequence 2, Appli
33 159.2 11.8 4411529 3 US-09-103-840A-1 Sequence 1, Appli
34 154 11.4 3333 3 US-09-532-803-10 Sequence 10, Appli
35 154 11.4 3333 4 US-09-653-403-1 Sequence 1, Appli
36 154 11.4 3333 4 US-10-013-784-1 Sequence 1, Appli
37 151 11.2 1364 3 US-09-095-855-204 Sequence 204, App
38 151 11.2 1364 4 US-09-205-426-204 Sequence 5, Appli
39 145 10.8 2375 2 US-08-931-800A-5 Sequence 931, App
40 144.2 10.7 3152 4 US-09-221-017B-931 Sequence 1, Appli
41 144 10.7 1830121 4 US-09-357-884-1 Sequence 1, Appli
42 144 10.7 1830121 4 US-09-643-990A-1 Sequence 1, Appli
43 142 10.5 2574 4 US-08-956-171E-441 Sequence 441, App
44 140.6 10.4 898 2 US-08-997-080-185 Sequence 185, App
45 140.6 10.4 898 2 US-08-997-362-185 Sequence 185, App

ALIGNMENTS

RESULT 1
US-09-878-766A-21
; Sequence 21, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer Gap4
; OTHER INFORMATION: chimeric GapC protein
; NAME/KEY: CDS
; LOCATION: (1)..(1347)
US-09-878-766A-21

Query Match 100.0%; Score 1347; DB 4; Length 1347;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAATAACAGGGATTATTTTCTTTGAGTCATTATTTCTGTCGATGC 60
Db 1 ATGAAAAAATAACAGGGATTATTTTCTTTGAGTCATTATTTCTGTCGATGC 60

QY 61 CAGGCAACTACGATCCGGTAGTAAAGTTGGTATTAAACGGTTTCGGTCGATC 120
Db 61 CAGGCAACTACGATCCGGTAGTAAAGTTGGTATTAAACGGTTTCGGTCGATC 120

QY 121 CGACGCTCTGATCCCGTCGATTCAAATGTTGAAGTGTGAAAGTAACTCGATCAAC 180
Db 121 CGACGCTCTGATCCCGTCGATTCAAATGTTGAAGTGTGAAAGTAACTCGATCAAC 180

QY 181 GACCTTACAGATCCAAACATGCTTGCACACTTTGTTGAATACGATPACAACTCAAGCGT 240
Db 181 GACCTTACAGATCCAAACATGCTTGCACACTTTGTTGAATACGATPACAACTCAAGCGT 240

QY 241 TTTCAGGCACTGTTGAAAGTAAAGAGTGGATTTGAAGTAAACGGAACCTTATCAAAA 300
Db 241 TTTCAGGCACTGTTGAAAGTAAAGAGTGGATTTGAAGTAAACGGAACCTTATCAAAA 300

QY	301	GT	TTCTGCTGAACG	TGATCCAGAAACATCGAC	ATGGGCAACTGACGGT	TTGAAATCGTT	360
Db	301	GT	TTCTGCTGAACG	TGATCCAGAAACATCGAC	ATGGGCAACTGACGGT	TTGAAATCGTT	360
QY	361	CT	GAAGCACTCGAGG	TACTGTAGAAAGTTAAAGAT	TGGTGGATTTCACGTT	TAAACGGAAAA	420
Db	361	CT	GAAGCACTCGAGG	TACTGTAGAAAGTTAAAGAT	TGGTGGATTTCACGTT	TAAACGGAAAA	420
QY	421	TT	CATTAAAGTTTCTGCT	GAAAAAGATCCAGAACAAATTCAT	CTGGGCAACTGACCGT	GTGT	480
Db	421	TT	CATTAAAGTTTCTGCT	GAAAAAGATCCAGAACAAATTCAT	CTGGGCAACTGACCGT	GTGT	480
QY	481	GAAATCGTTCTTTGAAAT	CGATGGTACTGTGTGAAGTTAAAGAAAGTGGATTCGAAAGT	TTAAAC	540		
Db	481	GAAATCGTTCTTTGAAAT	CGATGGTACTGTGTGAAGTTAAAGAAAGTGGATTCGAAAGT	TTAAAC	540		
QY	541	GGTCAATTTGTTAAAGTTTCTGCT	GAAACGCGACAGCAAACTGAT	CTGGGCTACTGAT	600		
Db	541	GGTCAATTTGTTAAAGTTTCTGCT	GAAACGCGACAGCAAACTGAT	CTGGGCTACTGAT	600		
QY	601	GGCGTAGAAATCGTTCTTTGAAGCAACTAGTTTCTTTTGCT	TAAAAAAGAGCTGCTGAAAAA	660			
Db	601	GGCGTAGAAATCGTTCTTTGAAGCAACTAGTTTCTTTTGCT	TAAAAAAGAGCTGCTGAAAAA	660			
QY	661	CACTTACATGCTAAACGGTGTCTTAAAGTTGTTATCA	CAGCTCTGGTGGAAAAACGCGTT	720			
Db	661	CACTTACATGCTAAACGGTGTCTTAAAGTTGTTATCA	CAGCTCTGGTGGAAAAACGCGTT	720			
QY	721	AAAACAGTTGTTTTCACACACTAACCAACACACACTTCTTGA	CGGTACTGTAACCAAGTTATCTCA	780			
Db	721	AAAACAGTTGTTTTCACACACTAACCAACACACACTTCTTGA	CGGTACTGTAACCAAGTTATCTCA	780			
QY	781	GGTGCTTCATGCTACTACAAACCTGTTTACGCTCCTATGGCT	TAAAGCTCTTCACGATGCAATTT	840			
Db	781	GGTGCTTCATGCTACTACAAACCTGTTTACGCTCCTATGGCT	TAAAGCTCTTCACGATGCAATTT	840			
QY	841	GGTATCCAAAAAGGTCCTTATGACTACAAATCCACGCTTATAC	TGTCGACCAATGATCGTT	900			
Db	841	GGTATCCAAAAAGGTCCTTATGACTACAAATCCACGCTTATAC	TGTCGACCAATGATCGTT	900			
QY	901	GACGGACACACCGTGGTGGTGACCTTCGTCGTGCTGTCGTGCTGTCGTGCTG	CGAAACATTTGTT	960			
Db	901	GACGGACACACCGTGGTGGTGACCTTCGTCGTGCTGTCGTGCTGTCGTGCTG	CGAAACATTTGTT	960			
QY	961	CCTAACTCAACTGTCGTCTGCTTAAAGCTTATCGGCTTGTATCC	CAGAAATGGAATGGTAAA	1020			
Db	961	CCTAACTCAACTGTCGTCTGCTTAAAGCTTATCGGCTTGTATCC	CAGAAATGGAATGGTAAA	1020			
QY	1021	CTTGATGGTGTGTCACCAACGTTTCTTGTTCCAACTGGATCAGTAA	CTGAGTTGGTGTGA	1080			
Db	1021	CTTGATGGTGTGTCACCAACGTTTCTTGTTCCAACTGGATCAGTAA	CTGAGTTGGTGTGA	1080			
QY	1081	ACTCTTGATAAAAAACGTTTCTGTTGACGAATCAACCGCTGCTATG	AAAGTGTCTTCAAC	1140			
Db	1081	ACTCTTGATAAAAAACGTTTCTGTTGACGAATCAACCGCTGCTATG	AAAGTGTCTTCAAC	1140			
QY	1141	GACAGTTTCGGTTACACTGAAATCCAAATGTTTCTTTCAGATATCG	TAGCGTGTCAATC	1200			
Db	1141	GACAGTTTCGGTTACACTGAAATCCAAATGTTTCTTTCAGATATCG	TAGCGTGTCAATC	1200			
QY	1201	GGTTCATTTGTTGACGCAACTCAAACTAAAGTTATGGAAGTTGAC	CGGATCACAAATGGTT	1260			
Db	1201	GGTTCATTTGTTGACGCAACTCAAACTAAAGTTATGGAAGTTGAC	CGGATCACAAATGGTT	1260			
QY	1261	AAAGTTGTATCATCGTATGACAAATGTCCTTACACTGCTCAACT	TGTTCTGTACACTT	1320			
Db	1261	AAAGTTGTATCATCGTATGACAAATGTCCTTACACTGCTCAACT	TGTTCTGTACACTT	1320			
QY	1321	GAGTATTTTGCAAAAATTCGCTAAATAA	1347				
Db	1321	GAGTATTTTGCAAAAATTCGCTAAATAA	1347				

```

RESULT 2
US-09-878-766A-11
; Sequence 11, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaines, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Streptococcus dysgalactiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) .. (1011)
US-09-878-766A-11

```

Query Match	60.6%;	Score 815.8;	DB 4;	Length 1011;
Best Local Similarity	97.4%;	MisMatch No. 8.1e-215;		
Matches 829; Conservative	0;	Indels 22;	Indels 0;	Gaps 0;
QY	497	TCGATGGTACTGTTTCAAGTTAAAGAGAGTGGATTTCGAGTTAAACGGTCAATTGTTGTTAAAG	556	
DB	161	TTGACGGNACHTTTGAAGTTTAAAGAGTGGATTTGAAGTAAACGGAACTTCATCAAG	220	
QY	557	TTTCTGTGAACGCGAACCCAGCAACATGTACTGGCTACTGATGGCGTAGAAATCGTTTC	616	
DB	221	TTTCTGTGAACGTGATCCAGAAACATCGACTGGCAACTGACGGTGTGAAATCGTTTC	280	
QY	617	TTGAAGCAACTAGTTTCTTCTCTAAAGAAAGCTGCTGAAAAACACTTACATGCTAACG	676	
DB	281	TGGAAGCAACTGGTTTCTTCTCTAAAGAAAGCTGCTGAAAAACACTTACATGCTAACG	340	
QY	677	GTGCTAAAAAGTTGTTTATCACAGCTCCTGGTGAACACGAGCTTAAACAGTGTGTTTCA	736	
DB	341	GTGCTAAAAAGTTGTTTATCACAGCTCCTGGTGAACACGAGCTTAAACAGTGTGTTTCA	400	
QY	737	ACACTAACACGACATTCCTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTACTA	796	
DB	401	ACACTAACACGACATTCCTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGTACTA	460	
QY	797	CAAACTGTTTAGTCTCTATGGCTAAAGCTCTTTCAGATGCATTTGGTATCCAAAAAGTTC	856	
DB	461	CAAACTGTTTAGTCTCTATGGCTAAAGCTCTTTCAGATGCATTTGGTATCCAAAAAGTTC	520	
QY	857	TTATGACTCAATCCAGCGTTTATCTGGTGACCAAAATGATCCTTTGACGGACACACCGTG	916	
DB	521	TTATGACTCAATCCAGCGTTTATCTGGTGACCAAAATGATCCTTTGACGGACACACCGTG	580	
QY	917	GTGGTAGCTTCGTGCGTGCTGCTGGTGCTGCAAACTTTGTTCTTAACCTCAACTGGTG	976	
DB	581	GTGGTAGCTTCGTGCGTGCTGCTGGTGCTGCAAACTTTGTTCTTAACCTCAACTGGTG	640	
QY	977	CTGCTAAAGCTATCGGTCCTTGTTATCCAGAAATTGAATGGTAACTTGATGGTGTCTGCAC	1036	
DB	641	CTGCTAAAGCTATCGGTCCTTGTTATCCAGAAATTGAATGGTAACTTGATGGTGTCTGCAC	700	
QY	1037	AACGTGTTCTGTTTCCAACTGGATCAGTAACTCAGTTGGTTGTAACTCTTGTATAAAAAAG	1096	
DB	701	AACGTGTTCTGTTTCCAACTGGATCAGTAACTCAGTTGGTTGTAACTCTTGTATAAAAAAG	760	
QY	1097	TTTCTGTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACACGACAGTTTCGGTTACA	1156	
DB	761	TTTCTGTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACACGACAGTTTCGGTTACA	820	
QY	1157	CTGAAGATCCAAATGTTTCTTTCAGATATCGTAGCGGTGTCATACGGTTTCAATGTTGAAC	1216	

Db 821 CTGAGATCAATGTTCTTCAGATATCGTAGCGTGTCATCGGTTCAATGTTGACG 880
QY 1217 CAACCTAACTAAAGTTATGGAAGTTGACGATCAAAATGGTTAAAGTTGATCATGTT 1276
Db 881 CAACCTAACTAAAGTTATGGAAGTTGACGATCAAAATGGTTAAAGTTGATCATGTT 940
QY 1277 ATGACATGAAGTCTTACACGCTCAACTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1336
Db 941 ATGACATGAAGTCTTACACGCTCAACTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1000
QY 1337 TCCTAAATAA 1347
Db 1001 TCCTAAATAA 1011

RESULT 3
US-07-928-462-1
; Sequence 1, Application US/07928462
; Patent No. 5328996
; GENERAL INFORMATION:
; APPLICANT: Boyle, Michael D.P.
; APPLICANT: Lottenberg, Richard C.
; APPLICANT: Broder, Christopher C.
; APPLICANT: von Mering, Gregory O.
; TITLE OF INVENTION: Bacterial Plasmid Receptors as
; TITLE OF INVENTION: Fibrinolytic Agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,462
; FILING DATE: 19920810
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/524,411
; FILING DATE: 16-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/330,849
; FILING DATE: 29-MAR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1125 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; STRAIN: M untypable
; INDIVIDUAL ISOLATE: 64/14
; IMMEDIATE SOURCE:
; CLONE: pRL015
; FEATURE:

; NAME/KEY: CDS
; LOCATION: 115..1122
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 115..1122
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /codon_start= 115
; OTHER INFORMATION: /function= "High-affinity binding of
; OTHER INFORMATION: /product= "Streptococcal plasmin receptor"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "plr"
; OTHER INFORMATION: /number= 1
; OTHER INFORMATION: /label= PLR
; US-07-928-462-1

Query Match 58.0%; Score 780.6; DB 1; Length 1125;
Best Local Similarity 94.8%; Pred. No. 4.2e-205;
Matches 807; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 497 TCGATGGTACTGTTGAAGTTAAAGAGGTGATTGCGAGTTAAAGGTCAATGTTGTTAAAG 556
Db 275 TTGATGGAACAGTTGAAGTTAAAGAGGTGATTGGAAGTAAACGGAACATTCATCAAG 334
QY 557 TTTCTGTGAACGGCAACCAACCAATGACTGGGTACTGTAGTGGGTAGAAATCGTTC 616
Db 335 TTTCTGTGAACGGTATCCAGAAACATCGACTGGCAACTGATGGGGTTGAAATCGTTC 394
QY 617 TTGAAGCAACTAGTTCTTTCTCTAAAGAAAGAGTCTGCTGAAAAACACTTACATGCTAACG 676
Db 395 TTGAAGCAACTAGTTCTTTCTCTAAAGAAAGAGTCTGCTGAAAAACACTTACATGCTAACG 454
QY 677 GTGCTAAAAAGTTGTTTATACACGCTCTCTGCTGAAACGAGCTTTAAACAGTTGTTTCA 736
Db 455 GTGCTAAAAAGTTGTTTATACACGCTCTCTGCTGAAACGAGTTTAAACAGTTGTTTCA 514
QY 737 ACACCTAACCCAGCATTTCTTCACGCTACTGAAACAGTTATCTCAGGTGCTTCATGACTA 796
Db 515 ACACCTAACCCAGCATTTCTTCACGCTACTGAAACAGTTATCTCAGGTGCTTCATGACTA 574
QY 797 CAACTGTTTAGTCTCTATGGCTTAAAGCTCTTCACGATGCTATTTGGTATCCAAAAAGGTC 856
Db 575 CAACTGTTTAGTCTCTATGGCTTAAAGCTCTTCACGATGCTATTTGGTATCCAAAAAGGTC 634
QY 857 TTATGACTAATCAATCAGCTTTATCTGCTGACCAAAATGATCTTTGACGGACCAACGGT 916
Db 635 TTATGACTAATCAATCAGCTTTATCTGCTGACCAAAATGATCTTTGACGGACCAACGGT 694
QY 917 GTGGTGACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 976
Db 695 GTGGTGACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 754
QY 977 CTGCTAAAGCTATCGGTCCTTGTATCCAGAAATGATGTTAACTTGTATGTTGCTGCTGAC 1036
Db 755 CTGCTAAAGCTATCGGTCCTTGTATCCAGAAATGATGTTAACTTGTATGTTGCTGCTGAC 814
QY 1037 AACGTGTTCTGTTCCCACTGGATCAGTAACTGAGTTGTTGTAACCTCTTCATATAAAACG 1096
Db 815 AACGTGTTCTGTTCCCACTGGATCAGTAACTGAGTTGTTGTAACCTCTTCATATAAAACG 874
QY 1097 TTTCTGTGACGAAATCAACGCTCTATGAAAGTGTCTTCAACGACAGTTTCGGTTACA 1156
Db 875 TTTCTGTGACGAAATCAACGCTCTATGAAAGTGTCTTCAACGACAGTTTCGGTTACA 934
QY 1157 CTGAGATCCAAATGTTTCTTCAGATATCGTAGCGGTGTCATAGGTTGTTGTTGACG 1216
Db 935 CTGAGATCCAAATGTTTCTTCAGATATCGTAGCGGTGTCATAGGTTGTTGTTGACG 994
QY 1217 CAACCTAACTAAAGTTATGGAAGTTGACGATCAAAATGGTTAAAGTTGATCATGTT 1276
Db 995 CAACCTAACTAAAGTTATGGAAGTTGACGATCAAAATGGTTAAAGTTGATCATGTT 1054
QY 1277 ATGACATGAAGTCTTACACGCTCAACTGTTGTTGCTGATGCTTGTGCTGCTTGTGCAAAA 1336


```

RESULT 5
US-09-878-766A-13
; Sequence 13, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Streptococcus agalactiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1011)
; US-09-878-766A-13

Query Match      52.1%; Score 702.2; DB 4; Length 1011;
Best Local Similarity 89.1%; Pred. No. 1.5e-183;
Matches 758; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 497 TCGATGGTACTGTTGAAGTTAAAGAGGTGGATTGCAAGTTAACGGTCAATTTGTTAAAG 556
DB 161 TCGACGGTACTGTTGAAGTTAAAGAGGTGGATTGCAAGTTAACGGTCAATTTGTTAAAG 220
QY 557 TTTCTGCTGAACGCGAACGAGCAAAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC 616
DB 221 TTTCTGCTGAACGCGAACGAGCAAAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC 280
QY 617 TTGAAGCAACTAGTTTCTTTGCTTAAAGAGAGCTGCTGAAACCACTTTACATGCTAACG 676
DB 281 TTGAAGCAACTAGTTTCTTTGCTTAAAGAGAGCTGCTGAAACCACTTTACATGCTAACG 340
QY 677 GTGCTTAAAGAGTTGTTATACAGCTCTGCTGGAACGCTTAAACAGTTGTTTCA 736
DB 341 GTGCTTAAAGAGTTGTTATACAGCTCTGCTGGAACGCTTAAACAGTTGTTTCA 400
QY 737 ACATTAACCAAGTATCTTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 796
DB 401 ACATTAACCAAGTATCTTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 460
QY 797 CAACCTGTTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 856
DB 461 CAACCTGTTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 520
QY 857 TTATGACTACAACTCCACGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 916
DB 521 TGAAGTACTATCCACGCTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580
QY 917 GTGGTGAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
DB 581 GTGGTGAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 640
QY 977 CTGCTTAAAGTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
DB 641 CTGCTTAAAGTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
QY 1037 AACGTGTTCTGTTCCAACTGGATCACTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1096
DB 701 AACGTGTTCTGTTCCAACTGGATCACTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 760
QY 1097 TTTCTGTTGACGAATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156
DB 761 TAACTGTGGAAGAGTAAATGCGAGCTATGAACGACGACGACGACGACGACGACGACGACGACGACG 820
QY 1157 CTGAAGATCCAAATGTTTCTTCATGATATCGTAGGCGGTGTCATACGGTTCATGTTTGGACG 1216

RESULT 6
US-09-878-766A-19
; Sequence 19, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Streptococcus iniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1011)
; US-09-878-766A-19

Query Match      50.3%; Score 678.2; DB 4; Length 1011;
Best Local Similarity 87.3%; Pred. No. 6e-177;
Matches 743; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 497 TCGATGGTACTGTTGAAGTTAAAGAGGTGGATTGCAAGTTAAGTGAAGTGAAGTGAAGTGAAG 556
DB 161 TTGACGGTACAGTTGAAAGTTAAAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 220
QY 557 TTTCTGCTGAACGCGAACGAGCAAAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC 616
DB 221 TTTCTGCTGAACGCGAACGAGCAAAACATTGACTGGGCTACTGATGGCGTAGAAATCGTTC 280
QY 617 TTGAAGCAACTAGTTTCTTTGCTTAAAGAGAGCTGCTGAAACCACTTTACATGCTAACG 676
DB 281 TTGAAGCAACTAGTTTCTTTGCTTAAAGAGAGCTGCTGAAACCACTTTACATGCTAACG 340
QY 677 GTGCTTAAAGAGTTGTTATACAGCTCTGCTGGAACGCTTAAACAGTTGTTTCA 736
DB 341 GTGCTTAAAGAGTTGTTATACAGCTCTGCTGGAACGCTTAAACAGTTGTTTCA 400
QY 737 ACATTAACCAAGTATCTTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 796
DB 401 ACATTAACCAAGTATCTTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 460
QY 797 CAACCTGTTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 856
DB 461 CAACCTGTTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 520
QY 857 TTATGACTACAACTCCACGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 916
DB 521 TGAAGTACTATCCACGCTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580
QY 917 GTGGTGAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
DB 581 GTGGTGAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 640
QY 977 CTGCTTAAAGTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
DB 641 CTGCTTAAAGTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
QY 1037 AACGTGTTCTGTTCCAACTGGATCACTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1096
DB 701 AACGTGTTCTGTTCCAACTGGATCACTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 760
QY 1097 TTTCTGTTGACGAATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156
DB 761 TAACTGTGGAAGAGTAAATGCGAGCTATGAACGACGACGACGACGACGACGACGACGACGACGACG 820
QY 1157 CTGAAGATCCAAATGTTTCTTCATGATATCGTAGGCGGTGTCATACGGTTCATGTTTGGACG 976
```



```

: APPLICANT: Fontaine, Michael
: TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
: TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
: FILE REFERENCE: 9000-0057
: CURRENT APPLICATION NUMBER: US/09/878.766A
: CURRENT FILING DATE: 2001-09-10
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 15
: LENGTH: 1011
: TYPE: DNA
: ORGANISM: Streptococcus uberis
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(1011)
: US-09-878-766A-15

```

QY	1277	ATGCAANTAGAAATGTCCTTACACTGCTCAACTTGTTCGTACACTTGAGTATTTTGCAGAAA	1333
Db	941	ATGACACGAGAAATGTCCTTACACTGCACAACTTGTTCGTACTCTTGTAGTACTTTGCAGAAA	1000
QY	1337	TCGCTAAATAA	1347
Db	1001	TCGCTAAATAA	1011
RESULT 9			
US-08-961-083-53			
; Sequence 53, Application US/08961083			
; Patent No. 6159469			
; GENERAL INFORMATION:			
; APPLICANT: Choi et. al.			
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines			
; NUMBER OF SEQUENCES: 452			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Human Genome Sciences, Inc.			
; STREET: 9410 Key West Avenue			
; CITY: Rockville			
; STATE: Maryland			
; COUNTRY: USA			
; ZIP: 20850			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage			
; COMPUTER: HP Vectra 486/33			
; OPERATING SYSTEM: MSDOS version 6.2			
; SOFTWARE: ASCII Text			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/961,083			
; FILING DATE:			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER:			
; FILING DATE:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Brookes, A. Anders			
; REGISTRATION NUMBER: 36,373			
; REFERENCE/DOCKET NUMBER: PB340P2			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (301) 309-8504			
; TELEFAX: (301) 309-8512			
; INFORMATION FOR SEQ ID NO: 53:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1000 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
US-08-961-083-53			
Query Match 49.0%; Score 660.2; DB 3; Length 1000;			
Best Local Similarity 87.9%; Pred: No. 5.3e-172;			
Matches 743; Conservative 0; Mismatches 98; Indels 4; Gaps 2;			
QY	497	TCGATGGTACTGTTGAAGTTAAAGAGGTGGATTCCGAAGTTAACGGTCAATTTGTTAAAG	556
Db	159	TCGACGGTACTGTTGAAGTTAAAGAGGTGGATTGAAGTTAACGGTAAATTCATCAAG	218
QY	557	TTTCTGCTGAACGCAACGACCAACATGTGCTGGCTACTGATGGCGTGAAGATCGTTC	616
Db	219	TTTCTGCTGAACGCGATCCAGACAAATCGACTGGCTACTGACGGGTAGAAAATCGTTC	278
QY	617	TTGAAGCAACTAGTTTCTTTGCTAAAAAGAGGTGCTGAAAAACACCTTACATGCTAACG	676
Db	279	TTGAAGCTACTGGTTTCTTTGCTAAGAGAGACGCTGAAAAACACCTTAAAGGT---	335
QY	677	GTGCTAAAAAGTTGTTATCATCAGCTCTCGTGGAAACGAGTTAAACAGTTGTTTCA	736
Db	336	GAGCTAAAAAGTTGTTATCATCTCTCTCGTGGAAACGAGTTAAACAGTTGTTTCA	395
QY	737	ACACTTAACCAACGACATTTCTTGACGGTACTGAAACAGATTATCTCAGGTGCTTCATGTACTA	796

Db 396 ACACATACACGAGCTTCTTGACGGTACTGAAACAGATTATCTCAGGTGCTTCACTACTA 455
Qy 797 CAAACTGTTAGCTTCTTAAAGCTCTTCCAGATGCAATTTGGTATCCAAAAGGTC 856
Db 456 CAAACTGTTGCTTCAATGCTAAAGCTCTTCAAGACAATTTGGTGTGTTGAAGAT 515
Qy 857 TTATGACTACATCCAGCTTATCTGTCACCAATGATCTTGAAGACCAACCGTG 916
Db 516 TGATGACTACTATCCAGCTTATCTGTCACCAATGATCTTGAAGACCAACCGTG 575
Qy 917 GTGTGACCTTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 576 GTGTGACCTTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 635
Qy 977 CTGCTAAGCTATCGCTTCTGTTATCCAGAAATGAATGTAATGATGCTGCTGCTGCTG 1036
Db 636 CTGCAAAAGCTATCGCTTCTGTTATCCAGAAATGAATGTAATGATGCTGCTGCTGCTG 695
Qy 1037 AACGTGTTCTGTTCCAACTGGATCAGTAACTGATGTTGTTGTAATCTTGTATAAAACG 1096
Db 696 AACGGTTTCCAACTGGATCAGTAACTGATGTTGTTGTAATCTTGTATAAAACG 755
Qy 1097 TTTCTGTTGACGAATCAACGCTGCTGTAAGAGCTGCTTCAACGACAGTTTCGGTTACA 1156
Db 756 TTTCTGTTGATGAGTGAACGCTGCTGTAAGAGCTGCTTCAACGACAGTTTCGGTTACA 815
Qy 1157 CTGAAGATCAATGTTTCTTCTGATATCTGTCGCTGCTGCTGCTGCTGCTGCTGCTG 1216
Db 816 CAGAGATCAATGCTATCTTCAAGATATCTGATGATGTTTACGTTTCAATGTTGACG 875
Qy 1217 CAACCTAACTAAAGTTATGGAAGTTGACGGATCACAATGTTGTTAAAGTTGATCATG 1276
Db 876 CAACCTAACTAAAGTTATGGAAGTTGACGGATCACAATGTTGTTAAAGTTGATCATG 935
Qy 1277 ATGACATGAATGCTTCTTACACTGCTCACTTCTTCTGCTGCTGCTGCTGCTGCTG 1335
Db 936 ACACACAGAAATGTCATACACTGCAACACTTCTGCTGCTGCTGCTGCTGCTGCTG 995
Qy 1336 ATGCG 1340
Db 996 ATTGC 1000

RESULT 10

US-09-536-784-53

; Sequence 53, Application US/09536784

; Patent No. 6573082

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION NUMBER:

; APPLICATION NUMBER: US/09/536,784

; FILING DATE: 30-Oct-1997

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/961,083

; FILING DATE: OCT-30-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Michelle S. Marks

; REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB340P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 1000 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-09-536-784-53

Query Match 49.0%; Score 660.2; DB 4; Length 1000;
Best Local Similarity 87.9%; Pred. No. 5.3e-172;
Matches 743; Conservative 0; Mismatches 98; Indels 4; Gaps 2;

Qy 497 TCGATGGTACTGTTGAAGTAAAGAGGTGATTGCAAGTTAAACGGTCAATTTGTTAAAG 556
Db 159 TCGACGGTACTGTTGAAGTTAAAGAGGTGATTGCAAGTTAAACGGTCAATTTGTTAAAG 218
Qy 557 TTTCTGCTGAACGGGAACCAACCAATTTGACTGGCTTACTGATGGGTAGAAATCGTTC 616
Db 219 TTTCTGCTGAACGGTATCCAGAACCAATCGACTGGCTTACTGATGGGTAGAAATCGTTC 278
Qy 617 TTGAAGCAACTAGTTTCTTCTGCTAAAGAGAGCTGCTGAAACACATTTACATGCTAACG 676
Db 279 TTGAAGCTACTGGTTTCTTCTGCTAAAGAGAGAGCTGCTGAAACACATTTACATGCTAACG 335
Qy 677 GTGCTAAAAAGTTGTTTATCATCAGCTCTCTGCTGAAACAGCGTTTAAACAGTTGTTTCA 736
Db 336 GAGCTAAAAAGTTGTTTATCATCAGCTCTCTGCTGAAACAGCGTTTAAACAGTTGTTTCA 395
Qy 737 ACATRACACGACATCTTTCAGCGTACTGAAACAGTTTCTCAGTGTCTCATGTACTA 796
Db 396 ACATRACACGACGTTTCTTTCAGCGTACTGAAACAGTTTCTCAGTGTCTCATGTACTA 455
Qy 797 CAAACTGTTTAGCTCCTATGCTTAAAGCTCTTTCACGATGCAATTTGGTATCCAAAAGGTC 856
Db 456 CAAACTGCTTGGCTTCCAAATGCTTAAAGCTCTTTCACGATGCAATTTGGTATCCAAAAGGTC 515
Qy 857 TTATGACTACATCCAGCTTATCTGCTGACCAATATGATCCTTGACGGACCAACCGTG 916
Db 516 TGATGACTACTATCCAGCTTATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 575
Qy 917 GTGTGACCTTTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 576 GTGTGACCTTTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 635
Qy 977 CTGCTAAGCTATCGCTTCTGTTATCCAGAAATGAATGTAATGATGCTGCTGCTGCTGCT 1036
Db 636 CTGCAAAAGCTATCGCTTCTGTTATCCAGAAATGAATGTAATGATGCTGCTGCTGCTGCT 695
Qy 1037 AACGTGTTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAATCTTGTATAAAACG 1096
Db 696 AACCGTTTCCAACTGGATCAGTAACTGAGTTGGTTGTAATCTTGTATAAAACG 755
Qy 1097 TTTCTGTTGACGAATCAACGCTGCTGTAAGAAAGTGTCTTCAACGACAGTTTCGGTTACA 1156
Db 756 TTTCTGTTGATGAGTGAACGCTGTAAGAAAGTGTCTTCAACGACAGTTTCGGTTACA 815
Qy 1157 CTGAAGATCAATTTGTTTCTTTCAGATATCTGAGCGTGTATACGGTTTCAATTTGTTGACG 1216
Db 816 CAGAGATCAATGCTATCTTTCAGATATCTGAGTATGTTTCAATTTGTTGACG 875
Qy 1217 CAACCTAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTTAAAGTTGATCATGTT 1276
Db 876 CAACCTAACTAAAGTTATGGAAGTTGACGGATCACAATTTGGTTAAAGTTGATCATGTT 935
Qy 1277 ATGCAATGAATGCTTCTTACACTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1335
Db 936 ACACACAGAAATGTCATACACTGCAACACTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 995

; Sequence 2676, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCES: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2676
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-2676

Query Match 31.7%; Score 426.8; DB 4; Length 1047;
Best Local Similarity 70.4%; Pred. No. 9.4e-108;
Matches 601; Conservative 0; Mismatches 247; Indels 6; Gaps 2;

Qy 497 TCGATGGTACTGTGTAAGTTAAAGAGGTGGATTGGAAGTTACGGTCAATTTGTTAAAG 556
Db |||||
Qy 197 TCACTGGAGAGTTGAAGTTATCGAAGTGGATTCCGTTGTAACGGTAAAGAAATTAAT 256
Db |||||
Qy 557 TTCTGCTGAACGACGACCAAGCAAAATTCAGTGGGCTACTGATGGGTAGAAATCGTTC 616
Db |||||
Qy 257 CATTCGATGACAGATGCTGTGTAAATTACCATGGGGCGGATTTAGATATCGAGTAGTAT 316
Db |||||
Qy 617 TTGAAGCAACTAGTTTCTTTGCTTAAAGAGCTGCTGTAAGAACTTACATGCTAAAG 676
Db |||||
Qy 317 TAGAATGCTACTGGTTCTTACTATCATATAAGAAAGCAAGCTCATCGATGC---AG 373
Db |||||
Qy 677 GTCTTAAAGAGTTGTTATCACAGCTCTGCTGTAAGCAAGCTTAAACAGTTGTTTCA 736
Db |||||
Qy 374 GTCTTAAAGAGTTAATTTCTAGTCCAGTAAAGTGATGTAAGAAACATCGTATTCA 433
Db |||||
Qy 737 ACCTAACCAAGCAATTTCTGACGGTACTGAACAGTTATCTCAGGTGCTTCTACTGTA 796
Db |||||
Qy 434 ACCTAACCAATGATACATTAAGTGTTCAGAAACAGTTGTTTCAAGTCTTCTTGTACTA 493
Db |||||
Qy 797 CAACCTGTTTACCTCTATGGCTTAAAGCTTTCAGCATGATTTGTTGTTTCCAAAGGTC 856
Db |||||
Qy 494 CTAACTCAATGACCAAGTTTGAAGAGTTTAAAGTACGAATTCGGTTTAGTTGAAGGTT 553
Db |||||
Qy 857 TTATGACTACAATCCAGCTTATCTGTTGACCAAAATGATCTTGTGACGACCAACCGTG 916
Db |||||
Qy 554 TCATGACTACAATTCAGCTTACACTGTTGACCAAAATACACAAGACGACCTCACAGAA 613
Db |||||
Qy 917 GTGGTACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
Db |||||
Qy 614 AAGGTGACAAACGTCGTGCAAGTGCAGCAGCTGAAATATATATCCCTAACTCAACAGGTG 673
Db |||||
Qy 977 CTGCTAAAGCTATCGGTCCTGTTTATCCAGAAATGGAATGTAACCTGATGGTGGTGCAC 1036
Db |||||
Qy 674 CTGCTAAAGCTATCGGTTAAGTTTATTCAGAAATCGATGTAATTTAGACGGTGGAGCAC 733
Db |||||
Qy 1037 AACGTGTTCTGTTCCAACTGGATACGATGATGATGATGATGATGATGATGATGATGATA---AAA 1093
Db |||||
Qy 734 AACGTGTTCCAGTTGCTACTGTTCTTCTTAACTGAATTAACGTGATGATGATGATGATGAT 793
Db |||||
Qy 1094 ACCTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTTTCGGTT 1153
Db |||||
Qy 794 ATGTAATGTTGACCAAGTTTAAAGTGTCTATGAAACAAAGCTTCTGACGATCAATTCGGTT 853
Db |||||
Qy 1154 ACCTGAAGATCAATTTGTTTCTGATATCGTAGCGGTGCTCATACGGTTCATTTGTTG 1213
Db |||||
Qy 854 ACCTGAAGACGAAATCGTATCTTCTGATATGTTGATGATGATGATGATGATGATGATGAT 913
Db |||||
Qy 1214 ACCTGAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1273
Db |||||

Db 914 ATGGACTCAACTCGTGTATGACTGTGGAGATCGTCAATTAGTTAAAGTTCAGCTT 973
Qy 1274 GGTATGACAAATGAATGTTTACACTGCTCAACTGTTTCTGACACTTGAGTATTTGCAA 1333
Db 974 GGTACGACAAATGAATGTTTACACTGCTCAACTGTTTCTGACACTTGAGTATTTGCTG 1033
Qy 1334 AAATCGCTAAATAA 1347
Db 1034 AACTTTCTAAATAA 1047

RESULT 15
US-08-956-171E-128
; Sequence 128, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 128:
SEQUENCE CHARACTERISTICS:
LENGTH: 6048 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 128:
US-08-956-171E-128

Query Match 30.4%; Score 409.2; DB 4; Length 6048;
Best Local Similarity 69.1%; Pred. No. 1.6e-102;
Matches 590; Conservative 0; Mismatches 258; Indels 6; Gaps 2;

Qy 497 TCGATGGTACTGTTGAAGTTAAAGAGGTGGATTGGAAGTTACGGTCAATTTGTTAAAG 556
Db 1352 TCACAGGTAAGTAGAGGTAGTTGATGGTGGTTTCCGCGTAAATGGTAAAGAGTTAAAT 1411
Qy 557 TTCTGCTGAACGCGCAACCAAGCAAAATTCAGTGGGCTACTGATGGGTAGAAATCGTTC 616
Db 1412 CATTCAGTGAACAGATGCAAGCAAAATTCACCTTCGAAAGACTTAAATATCGATGTAT 1471

Qy 617 TTGAGCACTAGTCTTCTTTGCTAAAGAAAGAGCTGCTGAAAGAAACACTTACATGCTAAG 676
Db |||||
1472 TAGAATGTACTGGTTTCTACACTGATAGATAGCAAGCTCATATTGAAGCA---G 1528
Qy 677 GTGCTAAAGAAAGTTGTTATCACAGCTCCTGCTGAAACGACGTTAAACAGTTGTTTCA 736
Db |||||
1529 GCGCTAAAGAAAGTATTAACTCTAGCACCAGCTACTGGTGACTTTAAAGAAACATCGTATTCA 1588
Qy 737 AACTAACCCAGGACATTCTTGAAGGTACTGAAGGTACTGAAGGTACTTCTAGGTGCTTCACTACTA 796
Db |||||
1589 AACTAACCCAGGACATTCTTGAAGGTACTGAAGGTACTGAAGGTACTTCTAGGTGCTTCACTACTA 1648
Qy 797 CAAACTGTTAGTCTTCTGCTAAAGCTCTTCAAGATGATTTGGTATCCAAAAAGGTC 856
Db |||||
1649 CAAACTGTTAGTCTTCTGCTAAAGCTCTTCAAGATGATTTGGTATCCAAAAAGGTT 1708
Qy 857 TTATGACTACATCCAGCTTACTGCTGACCAATGATCTTGAAGGTACTGAAGGTACTGAAGGT 916
Db |||||
1709 TAATGACTACATCCAGCTTACTGCTGACCAATGATCTTGAAGGTACTGAAGGTACTGAAGGT 1768
Qy 917 GTGCTGACCTTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db |||||
1769 AAGGTGACCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1828
Qy 977 CTGCTAAAGCTATCGGTCCTTGTATCCAGAAATGATGATGATGATGATGATGATGATGATGATG 1036
Db |||||
1829 CTGCTAAAGCTATCGGTCCTTGTATCCAGAAATGATGATGATGATGATGATGATGATGATGATG 1888
Qy 1037 AACGTGTTCTGTTCCCACTGATCAGTAACTGATGATGATGATGATGATGATGATGATGATG 1093
Db |||||
1889 AACGTGTTCTGTTCCCACTGATCAGTAACTGATGATGATGATGATGATGATGATGATGATG 1948
Qy 1094 ACCTTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAGCAGTTCGGTT 1153
Db |||||
1949 ACCTAACAGTTGAAAGTTAAAGAGCTATGAAAGCTGCTTCAAGCAGTTCGGTT 2008
Qy 1154 ACCTGAGATCCCAATTTGTTCTTCACTATCTAGGCTGCTATAGGTTCTATAGGTTCTATGTTG 1213
Db |||||
2009 ACCTGAGATCCCAATTTGTTCTTCACTATCTAGGCTGCTATAGGTTCTATAGGTTCTATGTTG 2068
Qy 1214 ACCTGAGATCCCAATTTGTTCTTCACTATCTAGGCTGCTATAGGTTCTATAGGTTCTATGTTG 1273
Db |||||
2069 ACCTGAGATCCCAATTTGTTCTTCACTATCTAGGCTGCTATAGGTTCTATAGGTTCTATGTTG 2128
Qy 1274 GGTATGACATGAAATGCTTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1333
Db |||||
2129 GGTATGACATGAAATGCTTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2188
Qy 1334 AAATCGCTAAATAA 1347
Db |||||
2189 AAATCGCTAAATAA 2202

Search completed: March 31, 2004, 03:10:34
Job time : 150 secs

Sequence 38002, A
Sequence 2, Appli
Sequence 15, Appli
Sequence 7, Appli
Sequence 53, Appli
Sequence 17, Appli
Sequence 9, Appli
Sequence 36070, A
Sequence 42211, A
Sequence 42029, A
Sequence 6750, Ap
Sequence 21110, A
Sequence 86, Appli
Sequence 9, Appli
Sequence 2058, Ap
Sequence 15335, A
Sequence 24308, A
Sequence 3778, Ap
Sequence 34706, A
Sequence 17082, A
Sequence 35550, A
Sequence 8442, Ap
Sequence 1563, Ap
Sequence 128, App
Sequence 128, App
Sequence 15, Appli
Sequence 21, Appli

16 678.2 50.3 1080 12 US-10-282-122A-38002
17 678.2 50.3 7571 12 US-10-158-844-2
18 676.6 50.2 1011 9 US-09-878-766A-15
19 676.6 50.2 1011 10 US-09-878-781-7
20 676.6 50.2 1011 14 US-10-134-297-7
21 660.2 49.0 1000 9 US-09-765-272-53
22 639.8 47.5 1011 9 US-09-878-766A-17
23 639.8 47.5 1011 14 US-09-878-781-9
24 639.8 47.5 1011 14 US-10-134-297-9
25 620.6 46.1 1011 12 US-10-282-122A-36070
26 567.8 42.2 1011 15 US-10-369-493-42211
27 481.4 35.7 1014 15 US-10-369-493-42029
28 479.6 35.6 1002 9 US-09-815-242-6750
29 478 35.5 999 12 US-10-282-122A-21110
30 478 35.5 4023 9 US-09-070-927A-86
31 454 33.7 684707 15 US-10-398-221-9
32 454 33.7 3011208 15 US-10-398-221-2058
33 453.6 33.7 1005 12 US-10-282-122A-15335
34 450.8 33.5 1011 12 US-10-282-122A-24308
35 447.6 33.2 4230 15 US-10-398-221-3778
36 423.8 31.5 1008 12 US-10-282-122A-34706
37 421.6 31.3 1005 12 US-10-282-122A-17082
38 419.4 31.1 1005 12 US-10-282-122A-35550
39 412.4 30.6 1011 9 US-09-815-242-8442
40 409.4 30.4 1008 9 US-09-815-242-4563
41 409.2 30.4 1011 12 US-10-282-122A-7813
42 409.2 30.4 6048 8 US-08-781-986A-128
43 409.2 30.4 6048 12 US-10-329-624-128
44 408.6 30.3 934 15 US-10-164-758-15
45 405.8 30.1 931 15 US-10-164-758-21

ALIGNMENTS

RESULT 1
US-09-878-766A-21
; Sequence 21, Application US/09878766A
; Patent No. US2002004928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: primer Gap4
; OTHER INFORMATION: chimeric GapC protein
; NAME/KEY: CDS
; LOCATION: (1)..(1347)
US-09-878-766A-21

Query Match 100.0%; Score 1347; DB 9; Length 1347;
Best Local Similarity 100.0%; Pred. No. 2.6e-308;
Matches 1347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAAATAACAGGATATTTTTCCTTCAGTCATATTCGTCTGCATGC 60
DB 1 ATGAAAAAATAACAGGATATTTTTCCTTCAGTCATATTCGTCTGCATGC 60
QY 61 CAGCAAACTACCGATCCGGTATCGGTAGTTAAAGTTGATTAACGGTTTCGGTCTGATC 120
DB 61 CAGCAAACTACCGATCCGGTATCGGTAGTTAAAGTTGATTAACGGTTTCGGTCTGATC 120

us-10-650-369-21.rnpb

OM nucleic - nucleic search, using sw model
Run on: March 31, 2004, 00:58:15 ; Search time 498 Seconds
(without alignments)
10070.068 Million cell updates/sec

Title: US-10-650-369-21
Perfect score: 1347
Sequence: 1 atgaaaaataacaggat.....ttgcaaaaatcgctaataa 1347

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1347	100.0	1347	9	US-09-878-766A-21
2	815.8	60.6	1011	9	US-09-878-766A-11
3	815.8	60.6	1011	10	US-09-878-781-3
4	815.8	60.6	1011	14	US-10-134-297-3
5	790.2	58.7	961	10	US-09-878-781-13
6	780.6	58.0	1011	12	US-10-282-122A-38195
7	778.2	57.8	1010	10	US-09-878-781-15
8	702.2	52.1	1011	9	US-09-878-766A-13
9	702.2	52.1	1011	10	US-09-878-781-5
10	702.2	52.1	1011	14	US-10-134-297-5
11	683	50.7	1080	9	US-09-815-242-9495
12	678.2	50.3	1011	9	US-09-878-766A-19
13	678.2	50.3	1011	10	US-09-878-781-11
14	678.2	50.3	1011	14	US-10-134-297-11
15	678.2	50.3	1080	9	US-09-815-242-9071

121 GGACGCTTCGATCGTGGTATCAAAATGTTGAAGGTTGCAAGTAACGTCATCAAC 180
121 GGACGCTTCGATCGTGGTATCAAAATGTTGAAGGTTGCAAGTAACGTCATCAAC 180
181 GACCTTACAGATCAAAACATGCTTGCACACTTGTGAAATACGATACAACTCAAGGACGT 240
181 GACCTTACAGATCAAAACATGCTTGCACACTTGTGAAATACGATACAACTCAAGGACGT 240
241 TTTGACGGAACCTGTTGAAGTTAAAGAGGTTGAAATGAAAGTAACCGGAACCTTCATCAA 300
241 TTTGACGGAACCTGTTGAAGTTAAAGAGGTTGAAATGAAAGTAACCGGAACCTTCATCAA 300
301 GTTCTCTGCTGAAGCTGATCAGAAACATCGACTGGGCAACTGACGGTGTGAAATCGTT 360
301 GTTCTCTGCTGAAGCTGATCAGAAACATCGACTGGGCAACTGACGGTGTGAAATCGTT 360
361 CTGGAAGCACTCGAGGGTACTGTAGAAAGTTAAAGATGTTGACGTTAAACGGAATA 420
361 CTGGAAGCACTCGAGGGTACTGTAGAAAGTTAAAGATGTTGACGTTAAACGGAATA 420
421 TTCAATTAAGTTCTGCTGAAAAGATCCAGAACTGACATGGGCAACTGACGGTGT 480
421 TTCAATTAAGTTCTGCTGAAAAGATCCAGAACTGACATGGGCAACTGACGGTGT 480
481 GAAATCGTTCTGAAATCGATGTTGTTGAAGTTAAAGAGGTTGCAAGTTAAAC 540
481 GAAATCGTTCTGAAATCGATGTTGTTGAAGTTAAAGAGGTTGCAAGTTAAAC 540
541 GGTCAATTTGTTAAAGTTCTGCTGAAAGCACTAGTTCTTTGCTAAAAAGAGGTCGTA 600
541 GGTCAATTTGTTAAAGTTCTGCTGAAAGCACTAGTTCTTTGCTAAAAAGAGGTCGTA 600
601 GGCCTAGAAATCGTTCTTGAAGCACTAGTTCTTTGCTAAAAAGAGGTCGTA 660
601 GGCCTAGAAATCGTTCTTGAAGCACTAGTTCTTTGCTAAAAAGAGGTCGTA 660
661 CACTTACATGCTGCTGAAAAGTTGTTATGACAGCTCTGCTGAAAAGAGGTCGTA 720
661 CACTTACATGCTGCTGAAAAGTTGTTATGACAGCTCTGCTGAAAAGAGGTCGTA 720
721 AAAACAGTTGTTTCAACACTAACACGACATCTTGAAGGTTGCTGAAAAGAGGTTCTCA 780
721 AAAACAGTTGTTTCAACACTAACACGACATCTTGAAGGTTGCTGAAAAGAGGTTCTCA 780
781 GGTGCTTCATGTAATAAAGTTGTTAGCTCTTATGAGGTTGCTGAAAAGAGGTTCTCA 840
781 GGTGCTTCATGTAATAAAGTTGTTAGCTCTTATGAGGTTGCTGAAAAGAGGTTCTCA 840
841 GGTATCCAAAAGGTTCTTATGACATCAATCCAGCTTATGCTGACCAAAATGATCCTT 900
841 GGTATCCAAAAGGTTCTTATGACATCAATCCAGCTTATGCTGACCAAAATGATCCTT 900
901 GACGACCAACCGTGGTGGTGAACCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 960
901 GACGACCAACCGTGGTGGTGGTGAACCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 960
961 CCTAATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
961 CCTAATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
1021 CTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
1021 CTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
1081 ACTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1081 ACTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1141 GACAGTTTCGGTTTACACTGAAAGTCCAAATGTTTCTTCCAGATATCGTAGGCGTGTATAC 1200
1141 GACAGTTTCGGTTTACACTGAAAGTCCAAATGTTTCTTCCAGATATCGTAGGCGTGTATAC 1200
1201 GGTTCATGTTTGACGCAACTCAAACTAAAGTTATGGAAGTTGACGATCACAATTCGTT 1260

RESULT 2

US-09-878-766A-11
; Sequence 11, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Streptococcus dysgalactiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1011)
US-09-878-766A-11

Query Match 60.6%; Score 815.8; DB 9; Length 1011;
Best Local Similarity 97.4%; Pred. No. 9.7e-183;
Matches 829; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
1201 GGTTCATGTTTGACGCAACTCAAACTAAAGTTATGGAAGTTGACGATCACAATTCGTT 1260
1261 AAAGTTGATCATGTTATGCAATGAAATGTCCTTACACTGCTCACTGCTCACTGTTGTTACACTT 1320
1261 AAAGTTGATCATGTTATGCAATGAAATGTCCTTACACTGCTCACTGCTCACTGTTGTTACACTT 1320
1321 GAGTATTTTGC AAAAATCGCTAAATAA 1347
1321 GAGTATTTTGC AAAAATCGCTAAATAA 1347
497 TCGATGTTACTGTTGAAAGTTAAAGAGTGAATTCGAAGTTAAAGGTTAAAGAGTGAATTCGAAG 556
161 TTGACGCAACTGTTGAAAGTTAAAGAGTGAATTCGAAGTTAAAGAGTGAATTCGAAG 220
557 TTTCTGCTGTAACCGCAACAGCAACATTTGACTGGGCTACTGATGGCGTAGAATTCGTTTC 616
221 TTTCTGCTGTAACCGTATCCAGAAACATCGACTGGCAACTGACGGTGTTCGAATTCGTTTC 280
617 TTGAAGCAACTAGTTTCTTTTGTCTAAAAAGAGTGTCTGAAAAACACTTACATCTACG 676
281 TGGAAAGCAACTGGTTTCTTTGCTAAAAAGAGTGTCTGAAAAACACTTACATCTACG 340
677 GTGCTAAAAAGTTGTTTATCACAGCTCTGCTGGAAGAGTGTCTGAAAAACAGTTGTTTTC 736
341 GTGCTAAAAAGTTGTTTATCACAGCTCTGCTGGAAGAGTGTCTGAAAAACAGTTGTTTTC 400
737 ACACCTAACCAACGACATCTTTGACGGTATCTGAACAGTTATCTCAGGTGCTTCATCTACTA 796
401 ACACCTAACCAACGACATCTTTGACGGTATCTGAACAGTTATCTCAGGTGCTTCATCTACTA 460
797 CAACTGTTTACTGCTTATGCTTAAAGTCTTTACAGATGCTTTTGGTATCTCCTAAAAAGGTC 856
461 CAACTGTTTACTGCTTATGCTTAAAGTCTTTACAGATGCTTTTGGTATCTCCTAAAAAGGTC 520
857 TTATGACTACAACTCAACGCTTATCTGTTGACCAAAATGATCTTACGAGGACCAACCGGTC 916
521 TTATGACTACAACTCAACGCTTATCTGTTGACCAAAATGATCTTACGAGGACCAACCGGTC 580
917 GTGGTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
581 GTGGTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 640
977 CTGCTAAAGCTATCGGTCTTGTGTTATCCCAAGATTAAGTGGTAAACTTGAATGGTGTGCTGAC 1036

Db 641 CTGCTAAAGCTATCGGTCCTTGTATCCAGAAATGGAATGGTAACTTTGATGGTGTGCAC 700
Qy 1037 AACGTGTTCTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACCTCTTGTATAAAACG 1096
Db 701 AACGTGTTCTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACCTCTTGTATAAAACG 760
Qy 1097 TTTCTGTTGACCAATCAACGCTGCTATGAAGCTGCTTCAACGACAGTTTCGGTTACA 1156
Db 761 TTTCTGTTGACCAATCAACGCTGCTATGAAGCTGCTTCAACGACAGTTTCGGTTACA 820
Qy 1157 CTGAAGATCCAAATGTTTCTTCAGATATCGTAGCGGTGTCATACGGTTTCATTGTTGACG 1216
Db 821 CTGAAGATCCAAATGTTTCTTCAGATATCGTAGCGGTGTCATACGGTTTCATTGTTGACG 880
Qy 1217 CAACTCAACTAAAGTTATGGAAGTTGACGATCAGAAATGGAATGGTAAAGTTGATCATGGT 1276
Db 881 CAACTCAACTAAAGTTATGGAAGTTGACGATCAGAAATGGAATGGTAAAGTTGATCATGGT 940
Qy 1277 ATGCAATGAATGCTTTACACTGCTCAACTTGTTCGTACACTTGGATTTTTCGCAAAA 1336
Db 941 ATGCAATGAATGCTTTACACTGCTCAACTTGTTCGTACACTTGGATTTTTCGCAAAA 1000
Qy 1337 TCGCTAAATAA 1347
Db 1001 TCGCTAAATAA 1011

RESULT 3
US-09-878-781-3
; Sequence 3, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; FILE OF INVENTION: STREPTOCOCCUS INFECTION
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Streptococcus dysgalactiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1011)
US-09-878-781-3

Query Match 60.6%; Score 815.8; DB 10; Length 1011;
Best Local Similarity 97.4%; Pred. No. 9.7e-183;
Matches 829; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 497 TCGATGCTACTGTTGAAGTTAAAGAGGTGATTCGAGTTAAAGTCAACGTCATTTGTTAAAG 556
Db 161 TTGACGGAATGTTGAAGTTAAAGAGGTGATTTGAAGTAAACGGAACCTTCATCAAG 220
Qy 557 TTTCTGCTGAACGCAACCAACGCAATTCAGTGGGCTACTGATGGCGTAGAAATCGTTC 616
Db 221 TTTCTGCTGAACGCTGATCCAGAAACATCGACTGGGCAACTGACGGTGGTAAATCGTTC 280
Qy 617 TTGAAGCACTAGTTCTTTGCTTAAAGAAAGCTGCTGAAACACTTACATGCTAAG 676
Db 281 TGAAGCACTAGTTCTTTGCTTAAAGAAAGCTGCTGAAACACTTACATGCTAAG 340
Qy 677 GTGCTAAAAAAGTTGTTATCACAGCTCCTGGTGAACACGCTTAAACAGTTGTTTCA 736
Db 341 GTGCTAAAAAAGTTGTTATCACAGCTCCTGGTGAACACGCTTAAACAGTTGTTTCA 400

Qy 737 AACTAAACCAAGCAATCTTGTGACGCTACTGAAACAGTTATCTCAGGTGCTTCATGCTACTA 796
Db 401 AACTAAACCAAGCAATCTTGTGACGCTACTGAAACAGTTATCTCAGGTGCTTCATGCTACTA 460
Qy 797 CAACTGTTTACTCTCTATGGCTAAAGCTCTTCAGATGCAATTTGGTATCCAAAAAGGTC 856
Db 461 CAACTGTTTACTCTCTATGGCTAAAGCTCTTCAGATGCAATTTGGTATCCAAAAAGGTC 520
Qy 857 TTATGACTACAACTCAGCGTTTATATCTGTGACCAAAATGATCTTTCAGCGACACACCGGTG 916
Db 521 TTATGACTACAACTCAGCGTTTATATCTGTGACCAAAATGATCTTTCAGCGACACACCGGTG 580
Qy 917 GTGGTGACCTTCGTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 581 GTGGTGACCTTCGTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
Qy 977 CTGCTAAAGCTATCGTCTTCTTATCCAGAAATGGAATGGTAAAGTTGATGCTGCTGCTGCTG 1036
Db 641 CTGCTAAAGCTATCGTCTTCTTATCCAGAAATGGAATGGTAAAGTTGATGCTGCTGCTGCTG 700
Qy 1037 AACGTGTTCTCTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACCTTTGATATAAAACG 1096
Db 701 AACGTGTTCTCTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACCTTTGATATAAAACG 760
Qy 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTTTTCGGTTACA 1156
Db 761 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAACGACAGTTTTCGGTTACA 820
Qy 1157 CTGAAGATCCAAATGTTTCTTCCAGATATCGTAGCGGTGTCATACGGTTTCATTGTTGACG 1216
Db 821 CTGAAGATCCAAATGTTTCTTCCAGATATCGTAGCGGTGTCATACGGTTTCATTGTTGACG 880
Qy 1217 CAACTCAAACTAAAGTTATGGAAGTTGACGATCACAATTTGGTTAAAGTTGATCATGGT 1276
Db 881 CAACTCAAACTAAAGTTATGGAAGTTGACGATCACAATTTGGTTAAAGTTGATCATGGT 940
Qy 1277 ATGCAATGAATGCTTTACACTGCTCAACTTGTTCGTACACTTGGATTTTTCGCAAAA 1336
Db 941 ATGCAATGAATGCTTTACACTGCTCAACTTGTTCGTACACTTGGATTTTTCGCAAAA 1000
Qy 1337 TCGCTAAATAA 1347
Db 1001 TCGCTAAATAA 1011

RESULT 4
US-10-134-297-3
; Sequence 3, Application US/10134297
; Publication No. US20030165524A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; FILE OF INVENTION: STREPTOCOCCUS INFECTION
; CURRENT APPLICATION NUMBER: US/10/134,297
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Streptococcus dysgalactiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1011)
US-10-134-297-3

Query Match 60.6%; Score 815.8; DB 14; Length 1011;
Best Local Similarity 97.4%; Pred. No. 9.7e-183;
Matches 829; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY	497	TGATGGTACTGTGAAGTTAAAGAAGTGATTGAAAGTTAACCGTCAATTTGTTAAAG	556
DB	161	TTGACGGAACTGTGAAGTTAAAGAAGTGATTTGAAGTAACAACGGAAACTTCATCAAAG	220
QY	557	TTTTCTGCTGAACCGAACCCAGCAAACTTGACTGGGCTACTGATGGCGTAGAAAATCGTTC	616
DB	221	TTTTCTGCTGAACCGTAGTCCAGAAAACTGACTGGGCAACTGACGGTGTGGAATCGTTC	280
QY	617	TTGAAGCAACTAGTCTTTCTTTGCTAAAAAGAAAGTCTGTA AAAACACTTATCATGCTAAACG	676
DB	281	TGGAAGCAACTGGTTCTTTCTTTGCTAAAAAGAAAGTCTGTA AAAACACTTATCATGCTAAACG	340
QY	677	GTCGTA AAAAAGTTGTTATCACAGCTCCTCGTGGAAACGACGTTAA AACAGTTGTTTCA	736
DB	341	GTCGTA AAAAAGTTGTTATCACAGCTCCTCGTGGAAACGACGTTAA AACAGTTGTTTCA	400
QY	737	ACACTAACCAAGCACTCTTGACGGTACTGA AACAGTTATCTCAGTGCTTCATGACTATA	796
DB	401	ACACTAACCAAGCACTCTTGACGGTACTGA AACAGTTATCTCAGTGCTTCATGACTATA	460
QY	797	CAAACTGTTTAGCTTCCTATGGCTAAAGCTCTT CACGATGCATTTGGTATCCAAAAAGGTC	856
DB	461	CAAACTGTTTAGCTTCCTATGGCTAAAGCTCTT CACGATGCATTTGGTATCCAAAAAGGTC	520
QY	857	TTATGACTACAAATCCACGCTTATCTGGTGACCA AATGATCCTTGACGGACCAACGGTG	916
DB	521	TTATGACTACAAATCCACGCTTATCTGGTGACCA AATGATCCTTGACGGACCAACGGTG	580
QY	917	GTCGTGACCTTCGTGCTGCTGCTGCTGCTGCTGCT GCAAACTTTGTTCTTAACCTCAACTGGTG	976
DB	581	GTCGTGACCTTCGTGCTGCTGCTGCTGCTGCTGCT GCAAACTTTGTTCTTAACCTCAACTGGTG	640
QY	977	CTGCTAAAGCTATCGGTCTTGTTATCCAGAA TTGAATGTAACTTGATGGTGTGTCAC	1036
DB	641	CTGCTAAAGCTATCGGTCTTGTTATCCAGAA TTGAATGTAACTTGATGGTGTGTCAC	700
QY	1037	AACGTGTTCTGTTTCCAACCTGGATCAGTA ACTGAGTTGGTTGTAACCTTTGATAAAAAACG	1096
DB	701	AACGTGTTCTGTTTCCAACCTGGATCAGTA ACTGAGTTGGTTGTAACCTTTGATAAAAAACG	760
QY	1097	TTTTCTGTTGACGAATCAACGCTGCTATGA AGTGCCTTCAACACGACGTTTCGGTTACA	1156
DB	761	TTTTCTGTTGACGAATCAACGCTGCTATGA AGTGCCTTCAACACGACGTTTCGGTTACA	820
QY	1157	CTGAAGATCAAATGTTTCTTTCAGATATCGT AGCGGTGTCATACGGTTCACTGTTTGACG	1216
DB	821	CTGAAGATCAAATGTTTCTTTCAGATATCGT AGCGGTGTCATACGGTTCACTGTTTGACG	880
QY	1217	CACCTCAAACTAAAGTTATGGAA GTTGACCGGATCA CAATTTGGTTAAAGTTGATCATGGT	1276
DB	881	CACCTCAAACTAAAGTTATGGAA GTTGACCGGATCA CAATTTGGTTAAAGTTGATCATGGT	940
QY	1277	ATGACAATGAAATGCTTTTACACTGCTCAACT GTTTCGTACACTTTGAGTATTTTGC AAAA	1336
DB	941	ATGACAATGAAATGCTTTTACACTGCTCAACT GTTTCGTACACTTTGAGTATTTTGC AAAA	1000
QY	1337	TCGCTAAATAA	1347
DB	1001	TCGCTAAATAA	1011

```

RESULT 5
US-09-878-781-13
; Sequence 13, Application US/09878781
; Publication No. US2003082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH gapC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION

```

```

; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 961
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: spyGapC
US-09-878-781-13

```

Query Match	58.7%	Score 790.2	DB 10	Length 961
Best Local Similarity	95.5%	Pred. No. 1.1e-178		
Matches 813	Conservative 0	Mismatches 36	Indels 0	Gaps 0
QY	497	TCGATGCTACTGTTGAAAGTTAAAGAAGGTGGATTCGAAGTTAACGGTCAATTTGTTAAAG	556	
Db	111	TGATGGAACAGTTTGAAGTTTAAAGAAGGTGGATTTGAAGTAAACGGAACTTCATCAAG	170	
QY	557	TTTCTGCTGAACGGACACAGCAAAACATTGACCTGGGCTACTGATGGCTAGAGAAATCGTTC	616	
Db	171	TTTCTGCTGAACGGATGATCCAGAAACATCGACTGGCAACTGATGGGTGAAATCGTTC	230	
QY	617	TGGAAGCAACTAGTTCCTTTGCTTAAAAAGAGAGCTGCTGAAAAACACTTACATGCTAACG	676	
Db	231	TGGAAGCAACTGGTTTCCTTGTCTAAAAAGAACAGCTGAAAAACACTTACATGCTAACG	290	
QY	677	GTGCTTAAAAAGTTGTTTATCACAGCTCTCTGGTGGAAACCAACGCTTAAAAACAGTTGTTTCA	736	
Db	291	GTGCTTAAAAAGTTGTTTATCACAGCTCTCTGGTGGAAACCAATGTTAAAAACAGTTGTTTCA	350	
QY	737	ACACTAACCCAGCAATTCCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTCATGTACTA	796	
Db	351	ACACTAACCCAGCAATTCCTTTGACGGTACTGAAACAGTTATCTCAGGTGCTCATGTACTA	410	
QY	797	CAAACTGTTTAGTCTCTATGGCTTAAAGTCTCTTACGATGCATTTGGTATCCAAAAAGTCT	856	
Db	411	CAAACTGTTTAGTCTCTATGGCTTAAAGTCTCTTACGATGCATTTGGTATCCAAAAAGTCT	470	
QY	857	TTATGACTACAACTCCAGCTTTTACTTGGTGTGACAAATGATCTCTGACGGACACACGGTG	916	
Db	471	TTATGACTACAACTCCAGCTTTTACTTGGTGTGACAAATGATCTCTGACGGACACACGGTG	530	
QY	917	GTGCTGACCTTCGTCGTGCTCGTCTGGTCTGCAAAACATTTCTCACTCAACTCGTGTG	976	
Db	531	GTGCTGACCTTCGTCGTGCAACGGCTGGTGTGCTGCTGCTTAACTCAACTCGTGTG	590	
QY	977	CTGCTTAAGCTATCGGTCTTGTTATCCAGAAATGAAATGGTTAACTTGATGGTGTGCTGAC	1036	
Db	591	CTGCTTAAGCTATCGGTCTTGTTATCCAGAAATGAAATGGTTAACTTGATGGTGTGCTGAC	650	
QY	1037	AACGTGTTCTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACTCTGATAAAAACG	1096	
Db	651	AACGTGTTCTCTGTTCCAACTGGATCAGTAACTGAGTTGGTTGTAACTCTGATAAAAACG	710	
QY	1097	TTTCTGTGTACGAAATCAACGCTGCTTATGAAGCTGCTTCAACACAGATTTCCGGTTACA	1156	
Db	711	TTTCTGTGTACGAAATCAACCTGCTATGAAGCTGCTTCAACACAGATTTCCGGTTACA	770	
QY	1157	CTGAAGATCCAAATGTTTCTTCCAGATATCGTAGGGCTGTCAATACGGTTCATTTGTTGACG	1216	
Db	771	CTGAAGATCCAAATGTTTCTTCCAGATATCGTAGGGCTGTCAATACGGTTCATTTGTTGACG	830	
QY	1217	CAACTCAAACTTAAGTTATGGAAGTTGACGGATCAAAATGGTTAAAGTGTATCATCGT	1276	
Db	831	CAACTCAAACTTAAGTTATGGAAGTTGACGGATCAAAATGGTTAAAGTGTATCATCGT	890	
QY	1277	ATGCAATGAATGTCTTACACTGCTCAACTGTTGTCGACACTTCAGTATTTTCCAAAA	1336	
Db	891	ATGCAACGAAATGTCTTACACTGCTCAACTGTTGTCGACACTTCAGTATTTTCCAAAA	950	

Qy 1337 TCGCTAAATAA 1347
 Db 1001 TCGCTAAATAA 1011
 RESULT 9
 US-09-878-781-5
 ; Sequence 5, Application US/09878781
 ; Publication No. US20030082781A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolton, Alexandra J.
 ; APPLICANT: Perez-Casal, Jose
 ; APPLICANT: Fontaine, Michael
 ; APPLICANT: Potter, Andrew A.
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
 ; FILE REFERENCE: 9000-0055
 ; CURRENT APPLICATION NUMBER: US/09/878,781
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 1011
 ; TYPE: DNA
 ; ORGANISM: Streptococcus agalactiae
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1011)
 US-09-878-781-5
 Query Match 52.1%; Score 702.2; DB 10; Length 1011;
 Best Local Similarity 89.1%; Pred. No. 7.1e-156;
 Matches 758; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
 Qy 497 TCGATGGTACTGTTGAAGTTAAGAGGTGATTCGAAGTTAAGCGTCAATTTGTTAAAG 556
 Db 161 TCGACGGTACTGTTGAAGTTAAGAGGTGATTCGAAGTTAAGCGTCAATTTGTTAAAG 220
 Qy 557 TTTCGCTGAACCGAACAGCAAAATTGCTGGCTACTGATGGCGTAGAAATCGTTC 616
 Db 221 TTTCGCTGAACCGAACAGCAAAATTGCTGGCTACTGATGGCGTAGAAATCGTTC 280
 Qy 617 TTGAACCACTAGTTTCTTTCGCTAAGAGGTGATTCGAAGTTAAGCGTCAATTTGTTTCA 736
 Db 341 GTGCTAATAAAGTTGTTATCACAGCTCCTGGTGGAAACGACGTTAAACAGAGTTGTTTCA 400
 Qy 737 ACACCTAACACAGCAATTCCTGAGCGTCTGAAACAGTTATCAGGTGCTTCATGTACTA 796
 Db 401 ACACCTAACACAGCAATTCCTGAGCGTCTGAAACAGTTATCAGGTGCTTCATGTACTA 460
 Qy 797 CAAACGTGTTTACTCTCTTCTGCTGAAGCTCTTCACGATGATTTGGTATCCAAAAGGTC 856
 Db 461 CAAACGTGTTTACTCTCTTCTGCTGAAGCTCTTCACGATGATTTGGTATCCAAAAGGTC 520
 Qy 857 TTATGACTACAAATCCACGCTTATCTGTTGACCAAAATGATCCTTTGACGGACCAACCGTG 916
 Db 521 TGATGACTACTATCCACGATACATCTGTTGACCAAAATGATCCTTTGACGGACCAACCGTG 580
 Qy 917 GTGGTGACCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
 Db 581 GTGGTGACCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
 Qy 977 CTGCTAAAGCTATCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
 Db 641 CTGCAAAAGCTATCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
 Qy 1037 AACGTGTTCTGCTTCCAACTGGATCACTGAGTTGGTTGTTAACTCTTGTATATAAAGC 1096


```
Db 461 CAAACTGCTCTGCTCCAAATGCTTAAGCTTTTACAAGACAACCTTTGGTGTAAACAAGGTT 520
Qy 857 TTATGACTACAAATCCAGCTTATATCTGCTGACCAAAATGATCCTTGACGACCAACACGGTG 916
Db 521 TGATGACTACTATCCAGCATACTAGTGGTACCAATGATCCTTGACGACCAACACGGTG 580
Qy 917 GTGTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 581 GTGTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
Qy 977 CTGCTAAAGCTATCCGCTCTGCTTATCCAGAAATGAAATGTAACCTTGATGCTGCTGCAC 1036
Db 641 CTGCAAAAGCTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 700
Qy 1037 AACGTGTTCTGCTTCCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1096
Db 701 AACGTGTTCTGCTTCCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 760
Qy 1097 TTTCTGTTGAGAAATCAACGCTGCTGATGAAAGCTGCTTCAAAAGCAGAGTTTCGGTTACA 1156
Db 761 TAACTGCTGAAAGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 820
Qy 1157 CTGAAGATCCAAATGCTTCTTTCAGATATCGTAGCGGTGTCATACGGTTCATGTTGTAAG 1216
Db 821 CTGAAGATCCAAATGCTTCTTTCAGATATCGTAGCGGTGTCATACGGTTCATGTTGTAAG 880
Qy 1217 CAATCAAACTAAAGTTATGGAAGTTGACGATGACGATGACGATGACGATGACGATGACGATG 1276
Db 881 CTACTCAAACTAAAGTTATGGAAGTTGACGATGACGATGACGATGACGATGACGATGACGATG 940
Qy 1277 ATGCAATGAATGCTTCTTACACTGCTCAACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1336
Db 941 ACGATACGAATGCTACACTTCACTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1000
Qy 1337 TCGCTAAATAA 1347
Db 1001 TCGCTAAATAA 1011
```

RESULT 11

```
US-09-815-242-9495
; Sequence 9495, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9495
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1080)
US-09-815-242-9495
```

Query Match

50.7%; Score 683; DB 9; Length 1080;
Best Local Similarity 88.5%; Pred. No. 2.5e-151;
Matches 753; Conservative 0; Mismatches 95; Indels 3; Gaps 1;

```
Qy 497 TCGATCGTACTGTTCAAGTTAAAGAGGTGATTCGAAGTTAAACGGTCAATTTGTTAAAG 556
Db 233 TCGACGGTACTGTTGAAGTTAAAGAGGTGATTCGAAGTTAAACGGTCAATTTGTTAAAG 292
Qy 557 TTTCTGCTGAAACCGCAACCCAGCAAACTTGAATGCTGGGCTACTGATGGCTAGAAATCGTTC 616
Db 293 TTTCTGCTGAAACCGTATCCAGAGCAAACTGCTGGGCTACTGATGGCTAGAAATCGTTC 352
Qy 617 TTGAAGCAACTAGTCTTCTTCTGCTAAAGAAAGCTGCTGAAAAACACTTACATGCTAAAG 676
Db 353 TTGAAGCTACTGTTTCTTCTGCTAAAGAAAGCTGCTGAAAAACACTTAAAGGT---G 409
Qy 677 GTGCTAAAGAAAGTCTTATACAGCTCTCTGCTGGAAGCAAGCTTAAACAGTTGTTTCA 736
Db 410 GAGCTAAAGAAAGTCTTATCACTGCTCTGCTGGAAGCAAGCTTAAACAGTTGTTTCA 469
Qy 737 ACATPAACACGACATCTTGAACGCTATGAAAAAGTATCTCAGGTGCTTCATGCTACTA 796
Db 470 ACATPAACACGACATCTTGAACGCTATGAAAAAGTATCTCAGGTGCTTCATGCTACTA 529
Qy 797 CAACTGTTAGTCTCTATGCTGCTAAAGCTCTTCAAGTGCATTTGGTATCCAAAAGGTC 856
Db 530 CAACTGCTTGGCTCCATGGCTAAAGCTCTTCAAGTGCATTTGGTATCCAAAAGGTC 589
Qy 857 TTATGACTACAATCAACGCTTATATCTGCTGACCAAAATGATCCTTGAACGACACACCGTG 916
Db 590 TGATGACTACTATCCAGCTTACACTGCTGACCAAAATGATCCTTGAACGACACACCGTG 649
Qy 917 GTGTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 650 GTGTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 709
Qy 977 CTGCTAAAGCTATCCGCTCTTGTATCCAGAAATGATGATGATGATGATGATGATGATGATGATG 1036
Db 710 CTGCAAAAGCTATCCGCTCTTGTATCCAGAAATGATGATGATGATGATGATGATGATGATGATG 769
Qy 1037 AACGTGTTCTGCTTCCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1096
Db 770 AACGTGTTCTGCTTCCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 829
Qy 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTCAAAAGCAGAGTTTCGGTTTACA 1156
Db 830 TTACTGTTGATGAGAGTGAACGAGCTATGAAAGCAGCTTCAACGAAATCATACGCTTACA 889
Qy 1157 CTGAAGATCCAAATGCTTCTTCAAGATATCGTAGCGGTGTCATACGGTTCATGTTGACG 1216
Db 890 CAGAAGATCCAAATGCTTCTTCAAGATATCGTAGCGGTGTCATACGGTTCATGTTGACG 949
Qy 1217 CAACTCAAACTAAAGTTATGGAAGTTGACGATCAAAATGCTTAAAGTTGATCATCGGT 1276
Db 950 CAACTCAAACTAAAGTTATGGAAGTTGACGATCAAAATGCTTAAAGTTGATCATCGGT 1009
Qy 1277 ATGACAATGAATGCTTACACTGCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1336
Db 1010 ACGACAACGAAATGCTACACTGCAAACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1069
Qy 1337 TCGCTAAATAA 1347
Db 1070 TCGCTAAATAA 1080
```


RESULT 12

US-09-878-766A-19
; Sequence 19, Application US/09878766A

; Patent No. US20020044928A1

; GENERAL INFORMATION:

; APPLICANT: Potter, Andrew A.

; APPLICANT: Perez-Casal, Jose

; APPLICANT: Fontaine, Michael

; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN

; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION

; FILE REFERENCE: 9000-0057

; CURRENT APPLICATION NUMBER: US/09/878,766A

; CURRENT FILING DATE: 2001-09-10

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 19

; LENGTH: 1011

; TYPE: DNA

; ORGANISM: Streptococcus iniae

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1011)

US-09-878-766A-19

Query Match 50.3%; Score 678.2; DB 9; Length 1011;

Best Local Similarity 87.3%; Pred. No. 3.4e-150;

Matches 743; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 497 TCGATCGTACTGTTGAAGTTAAAGAGGTGGATTGCGAAGTTAACGGTCAATTTGTTAAAG 556

DB 161 TTGACGGTACAGTTGAAGTTAAAGAGGTGGATTGCGAAGTTAACGGTCAATTTGTTAAAG 220

QY 557 TTCTCTGTAACCGAAGCAGCAACCAATTCAGTGGGCTACTGATGGCTAGAGAAATCGTTC 616

DB 221 TTCTCTGTAACCGAAGCAGCAACCAATTCAGTGGGCTACTGATGGCTAGAGAAATCGTTC 280

QY 617 TTGAAGCAACAGTCTTCTTTGCTTAAAGAGAGAGCTGCTGAAAAACACTTACATGCTAACG 676

DB 281 TTGAAGCAACAGTCTTCTTTGCTTAAAGAGAGAGCTGCTGAAAAACACTTACATGCTAACG 340

QY 677 GTGCTAAAGAGTTGTTATACAGCTCTCTGGTGAACAGCTTAAACAGTGTGTTTCA 736

DB 341 GTGCGAAAAAGTTGTTATACAGCTCTCTGGTGAACAGCTTAAACAGTGTGTTTCA 400

QY 737 ACATAACACAGCAATTTTGACGGTACTGAAACAGTATCTCAGGTGCTTTCATGTACTA 796

DB 401 ACATAACCAATGATTTCTTGATGAACAGCAATTTTCAGGTGCTTTCATGTACTA 460

QY 797 CAAACTGTTAGCTCTCTAGCTTAAAGCTCTTACAGTGAATTTGATTCCTAAACAGGTT 856

DB 461 CAAACTGTTAGCTCTCTAGCTTAAAGCTCTTACAGTGAATTTGATTCCTAAACAGGTT 520

QY 857 TTATGACTACAATCCAGCTTACTGTTGACCAAAATGATCTTGAACGACACACCGTG 916

DB 521 TAATGACTACTATCCATGTTTACCTGTTGACCAAAATGATCTTGAACGACACACCGTG 580

QY 917 GTGGTGAACCTTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976

DB 581 GTGGTGAACCTTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640

QY 977 CTGCTAAAGCTATCGGCTGTTGTTATCCAGAAATGATGTAACCTTGAATGTTGCTGCAC 1036

DB 641 CTGCTAAAGCAATCGGCTGTTGTTATCCAGAAATGATGTAACCTTGAATGTTGCTGCAC 700

QY 1037 AAGCTGTTCTGTTTCCAACTGGATCAGTAACTGAGTTGGTTGTTGTTGTTGTTGTTGTTG 1096

DB 701 AAGCTGTTCTGTTTCCAACTGGATCAGTAACTGAGTTGGTTGTTGTTGTTGTTGTTGTTG 760

QY 1097 TTCTGTTGAGAAATCAAGCTGCTGTTGTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1156

DB 761 CTTTCAAGTAAAGAAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 820

QY 1157 CTGAAGATCCAATTTGTTTCTTTCAGATATCGTAGCGTGTCTATAGGTTTCAATTTGTTCAACG 1216

DB 821 CTGAAGATGCTATGCTGTATCATCATGATATGATGTTTCTTACGGTTTCATTTATTTGATG 880

QY 1217 CAACTCAAACTTAAAGTTATGGAAGTTTACCGGATCACAATTTGGTTAAAGTTTGTATCATGGT 1276

DB 881 CTACTCAAACTTAAAGTTATGGAAGTTTACCGGATCACAATTTGGTTAAAGTTTGTATCATGGT 940

QY 1277 ATGCAATGAAATGCTTTCATCTCACTGCTCAACTTGTTCGTACACTTGTAGTATTTGCAAAAA 1336

DB 941 ATGCAATGAAATGCTTTCATCTCACTGCTCAACTTGTTCGTACACTTGTAGTATTTGCAAAAA 1000

QY 1337 TCGCTAAATAA 1347

DB 1001 TCGCTAAATAA 1011

RESULT 13

US-09-878-781-11

; Sequence 11, Application US/09878781

; Publication No. US20030082781A1

; GENERAL INFORMATION:

; APPLICANT: Bolton, Alexandra J.

; APPLICANT: Perez-Casal, Jose

; APPLICANT: Fontaine, Michael

; APPLICANT: Potter, Andrew A.

; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST

; TITLE OF INVENTION: STREPTOCOCCUS INFECTION

; FILE REFERENCE: 9000-0055

; CURRENT APPLICATION NUMBER: US/09/878,781

; CURRENT FILING DATE: 2002-09-10

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 11

; LENGTH: 1011

; TYPE: DNA

; ORGANISM: Streptococcus iniae

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1011)

US-09-878-781-11

Query Match 50.3%; Score 678.2; DB 10; Length 1011;

Best Local Similarity 87.3%; Pred. No. 3.4e-150;

Matches 743; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 497 TCGATGCTACTGTTTGAAGTTAAAGAGGTGGATTGCGAAGTTTAAACGGTCAATTTGTTAAAG 556

DB 161 TTGACGGTACAGTTGAAGTTAAAGAGGTGGATTGCGAAGTTTAAACGGTCAATTTGTTAAAG 220

QY 557 TTCTCTGTAACCGAAGCAGCAACCAATTCAGTGGGCTACTGATGGCTAGAGAAATCGTTC 616

DB 221 TTCTCTGTAACCGAAGCAGCAACCAATTCAGTGGGCTACTGATGGCTAGAGAAATCGTTC 280

QY 617 TTGAAGCAACAGTCTTCTTTGCTTAAAGAGAGAGCTGCTGAAAAACACTTACATGCTAACG 676

DB 281 TTGAAGCAACAGTCTTCTTTGCTTAAAGAGAGAGCTGCTGAAAAACACTTACATGCTAACG 340

QY 677 GTGCTAAAGAGTTGTTATACAGCTCTCTGGTGAACAGCTTAAACAGTGTGTTTCA 736

DB 341 GTGCGAAAAAGTTGTTATACAGCTCTCTGGTGAACAGCTTAAACAGTGTGTTTCA 400

QY 737 ACATAACCAAGCAATTTTCAGCGTACTGAAACAGTATCTCAGGTGCTTTCATGTACTA 796

DB 401 ACATAACCAAGCAATTTTCAGCGTACTGAAACAGTATCTCAGGTGCTTTCATGTACTA 460

QY 797 CAAACTGTTAGCTCTCTATGCTTAAAGCTCTTTCAGTATGCTATTTGTTGTTTCAAAAGGTC 856

DB 461 CAAACTGTTAGCTCTCTATGCTTAAAGCTCTTTCAGTATGCTATTTGTTGTTTCAAAAGGTC 520

QY 857 TTATGACTACAATCCAGCTTATCTGCTGTAACCAATGATCTTTCAGCGACCAACCGTG 916

1217 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTGGTTAAAGTTGTATCATGGT 1276
1009 CAACCTCAAACTAAAGTTATGGAAGTTGACGGATCACAATTGGTTAAAGTTGTATCATGGT 1009
1277 ATGACATGAAATGCTTCTTACACTGCTCAACTTGTTCGTACACTTGTAGTATTTTCBAAA 1336
1010 ACGACACGAAATGCTTACACTGCTCAACTTGTTCGTACACTTGTAGTATTTTCBAAA 1069
1337 TCGCTAAATAA 1347
1070 TTGCTAAATAA 1080

Search completed: March 31, 2004, 03:18:54
Job time : 509 secs

PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9071
LENGTH: 1080
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1080)
US-09-815-242-9071

Query Match 50.3%; Score 678.2; DB 9; Length 1080;
Best Local Similarity 88.1%; Pred. No. 3.5e-150;
Matches 750; Conservative 0; Mismatches 98; Indels 3; Gaps 1;

QY 497 TCGATGGTACTGTGGAAGTTAAAGAGGTGGAATTCGAAGTTAAGGTTACGGTCAATTGTTAAAG 556
Db 233 TCGACGGTACTGTGGAAGTTAAAGAGGTGGAATTCGAAGTTAAGGTTAATTCATCAAG 292
QY 557 TTCTCTGTGACCGACGACGACCAATTCAGTGGCTACTGATGGCTAGGAATCGTTC 616
Db 293 TTCTCTGTGACCGACGACGACCAATTCAGTGGCTACTGATGGCTAGGAATCGTTC 352
QY 617 TTGAAGCAACTAGTCTTCTTGTCTTAAAGAGGCTGCTGAAACCACTTACATGCTAACG 676
Db 353 TTGAGCTACTGTGTTCTTCTTGTCTTAAAGAGGCTGCTGAAACCACTTAAAGGT---G 409
QY 677 GTGCTTAAAGAGTTGTTATCAGAGCTCTGCTGGTGGAAACGAGCTTAAACAGTTGTTTCA 736
Db 410 GAGCTTAAAGAGTTGTTATCAGAGCTCTGCTGGTGGAAACGAGCTTAAACAGTTGTTTCA 469
QY 737 ACCTTACACGACATCTTCTGAGGCTACTGAAACAGTTATCTCAGGTGCTTCTATGTA 796
Db 470 ACCTTACACGACGTTCTTCTGAGGCTACTGAAACAGTTATCTCAGGTGCTTCTATGTA 529
QY 797 CAACTGTTTGTCTCTTAAAGGCTCTTCAAGATGATTTGGTATCCAAAAGGTC 856
Db 530 CAACTGTTTGTCTCTTAAAGGCTCTTCAAGATGATTTGGTATCCAAAAGGAT 589
QY 857 TTATGACTACATCCAGCTTATCTGCTGACCAATGATCCTTGGACGACACACCGTG 916
Db 590 TGATGACTATCTCCAGCTTATCTGCTGACCAATGATCCTTGGACGACACACCGTG 649
QY 917 GTGATGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db 650 GTGATGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 709
QY 977 CTGCTTAAAGCTATCGGCTTGTATCCAGATGATGATGATGATGATGATGATGATGATGATG 1036
Db 710 CTGCTTAAAGCTATCGGCTTGTATCCAGATGATGATGATGATGATGATGATGATGATGATG 769
QY 1037 AACGTGTTCTGTTCCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1096
Db 770 AACGTGTTCTGTTCCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 829
QY 1097 TTCTGTTGACGAATCAAGCTGCTATGAAAGCTGCTTCAAGAGGCTTTCGGTTTACA 1156
Db 830 TTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 889
QY 1157 CTGAAGATCAATGTTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1216
Db 890 CAGAGATCAATGTTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 949

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2004, 22:21:00 ; Search time 3333 Seconds

Title: US-10-650-369-21
Perfect score: 1347
Sequence: 1 atgcaaaaaataacagggat.....ttgcataaatcgctaaataa 1347

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

2068.514 Million cell updates/sec

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

ALIGNMENTS

RESULT 1

BH770540

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BH770540

LMGtag303

MG1363

Random Sequence

Tag Library

Lactococcus lactis

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

BH770540

LMGtag303

MG1363

Random Sequence

Tag Library

Lactococcus lactis

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

BH770540

LMGtag303

MG1363

Random Sequence

Tag Library

Lactococcus lactis

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

BH770540

LMGtag303

MG1363

Random Sequence

Tag Library

Lactococcus lactis

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

BH770540

LMGtag303

MG1363

Random Sequence

Tag Library

Lactococcus lactis

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

BH770540

LMGtag303

MG1363

Random Sequence

Tag Library

Lactococcus lactis

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.

1127 bp

DNA

linear

GSS 01-MAY-2002

subsp. cremoris

genomic, genomic survey sequence.


```

FEATURES
  source      Location/Qualifiers
    1..1127
    /organism="Lactococcus lactis subsp. cremoris"
    /mol_type="genomic DNA"
    /strain="MG1363"
    /sub_species="cremoris"
    /db_xref="taxon:1359"
    /clone_lib="MG1363 Random Sequence Tag Library"
    /notes="Vector: pSGM2; Site 1: SmaI; Library of
    chromosomal fragments of L. lactis strain MG1363 was
    prepared by partial AluI digestion or by sonication."

ORIGIN
  Query Match      35.0%; Score 471.8; DB 28; Length 1127;
  Best Local Similarity 73.0%; Pred. No. 1.1e-110;
  Matches 621; Conservative 0; Mismatches 227; Indels 3; Gaps 1;

Qy 500 ATGCTACTGTTGAAGTTAAAGAGGTTGATTCGAGTTACGCTCAATTTGTTAAAGTTT 559
Db 153 AAGAAACAGTTGAGTCAAGAGAGATGGTTTGTATGTCACGGAATTTGTGAAGTCA 212
Qy 560 CTGCTGAACGCGAACACGACCAATTCAGTGGCTACTGATGCGTAGAAATCGTTCTTG 619
Db 213 CTGCTGAACGTAATCCAGAGATATTCATGGCTGATTCGTGTGGAAATTTGTTCTTG 272
Qy 620 AAGCAACTAGTTCTTTGCTTAAAGAAAGAGCTGCTGAAACACACTTACATGCTAACGGTG 679
Db 273 AAGCGACAGGTTCTTTGCGCATTAAGAAAAGCTGAAAACACTTTGCACTCTGGTGGAG 732
Qy 680 CTAAAAAGTTGTTATCACAGCTCTGCTGGAACGACGTTAAAAAGAGTTGTTTCAACA 739
Db 333 CTAAGAAAGTTTGTATTACAGCACCTGCTGGAATGATGTCAAAACAGTCGTTATCAACA 799
Qy 740 CTAACCAACGACATTTCTGACGCTACTGAAACAGTTATCTCAGGTGCTTCTATGTACTACAA 799
Db 393 CAATATCACAAATTTCTGATGAGTGAAGTGAACAGTTATTTCTGCGGTTCTTTGTACTACAA 452
Qy 800 ACTGTTTGTGCTTATCGCTTAAAGCTCTTCCAGATGCTTTGTTATCAAAAAGGCTTA 859
Db 453 ATAGTTTGAACCAATGGCAGACGCAATTAACAAAACCTTTGTTGTTAAAGTTGGACAA 512
Qy 860 TGACTACAATCCACGCTTATCTGTCACCAATATGATCTCTGACGACGACCAACCGTGGTG 919
Db 513 TGACTACTGTCCTCATACATCTGTCACCAATATGATCTCTGACGACGACCAACCGTGGTG 572
Qy 920 GTGACCTTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 979
Db 573 GAGATTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 632
Qy 980 CTAAGAGCTATCGGCTTTGTTATCCAGATTTGAATGTTAAACTTTGATGCTGCTGCAAC 1039
Db 633 CTAAGCAATTTGCTTTGTTTCTGCTGCTGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 692
Qy 1040 GTGTTCTGTTTCCAACTGGATCAGTAACTGATGTTGTTGTTGTTGTTGTTGTTGTTGTT 1099
Db 693 GTGTTTCAACTCCAAACAGGATCAATTAAGTGAATTTGTTGATGCTGTTGAAACATGTCA 752
Qy 1100 CTGTTTGAAGAAATCAACGCTGCTATGAAAGCTGCTTCAAAACGACAGTTTCGTTTACACTG 1159
Db 753 CTGTTTGAAGAAATCAACGACGATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 812
Qy 1160 AAGATCCAAATTTGTTTTCAGATATCGTAGCGGTGTCATACGTTTCAATGTTTGAACGCA 1219
Db 813 TTGATGAAATTTGTTTTCATCTGACATTTATCGCATGCTGCTGCTGCTGCTGCTGCTGCTG 872
Qy 1220 CTCAACTTAAAGTTATGGA---AGTTGACGATCACAAATTTGTTTAAAGTTGTTATCATGTT 1276
Db 873 CATTGACTGAAGTGAACAGACTTAAAGACGCTGACAAATTAGTTTAAACTGCTGCTTGGT 932
Qy 1277 ATGCAATGAATGCTTTTACATGCTCACTGCTCACTGTTGTTGTTGTTGTTGTTGTTGTTG 1336
Db 933 ATGCAATGAATGAGTTTCACTGCTCAATTAATTCGTTACCTTGAATATTTTGGCCAAA 992

```

```

Qy 1337 TCGCTAAATAA 1347
Db 993 TTGCTAAATAA 1003

RESULT 2
LOCUS      CB686102
DEFINITION Bn01b_02m19 A
            990 bp mRNA linear EST 09-APR-2003
            Bn01b_AAPC_ECORC transgenic Brassica napus overexpressing BNCB17_c
            onstitutively_frost_tolerant Brassica napus cDNA clone Bn01b_02m19,
            mRNA sequence.
ACCESSION  CB686102
VERSION    CB686102.1 GI:29689827
KEYWORDS  EST.
SOURCE     Brassica napus (rape)
ORGANISM  Brassica napus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE  1 (bases 1 to 990)
            Singh, J., Allard, G., Tinker, N., Robert, L., Lacroix, C., De Moors, A.,
            Chagnon, J., Farah, S., Couroux, P. and Hattori, J.
            Expressed Sequence Tags from constitutively frost tolerant
            transgenic Brassica napus overexpressing BNCB17
            Unpublished (2002)
            Contact: Singh, J.A.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA
            OC6, Canada
            Tel: (613) 759-1662
            Fax: (613) 759-1701
            Email: singhja@em.agr.ca.
            Location/Qualifiers
              1..990
              /organism="Brassica napus"
              /mol_type="mRNA"
              /cultivar="Westar"
              /db_xref="taxon:3708"
              /clone="Bn01b_02m19"
              /tissue_type="fourth leaf"
              /dev_stage="3 weeks seedling grown at room temperature"
              /clone_lib="Bn01b_AAPC_ECORC transgenic Brassica napus_ove
              repressing BNCB17 constitutively frost tolerant"
              /note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI;
              Site 2: XhoI; Germinated in soil flats and seedlings grown
              for 3 weeks in a Coviron B-15 cabinet set at 20oC /16 hr
              light (250 Em-2sec-1) and 16 oC / 8 hr dark. Fourth leaves
              collected at 9 am and immediately frozen."

FEATURES
  source      Location/Qualifiers
    1..990
    /organism="Brassica napus"
    /mol_type="mRNA"
    /cultivar="Westar"
    /db_xref="taxon:3708"
    /clone="Bn01b_02m19"
    /tissue_type="fourth leaf"
    /dev_stage="3 weeks seedling grown at room temperature"
    /clone_lib="Bn01b_AAPC_ECORC transgenic Brassica napus_ove
    repressing BNCB17 constitutively frost tolerant"
    /note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI;
    Site 2: XhoI; Germinated in soil flats and seedlings grown
    for 3 weeks in a Coviron B-15 cabinet set at 20oC /16 hr
    light (250 Em-2sec-1) and 16 oC / 8 hr dark. Fourth leaves
    collected at 9 am and immediately frozen."

ORIGIN
  Query Match      14.5%; Score 195.6; DB 14; Length 990;
  Best Local Similarity 54.8%; Pred. No. 2.5e-39;
  Matches 456; Conservative 0; Mismatches 364; Indels 12; Gaps 3;

Qy 519 AGAAGGTGATTCGAAGTTAAGTTGTTTAAAGTTTCTGCTGACCGACCGACG 578
Db 132 AGACTCTGCTCTCTCCGTTGATGGAAGATCATCAAGATTGCTCTGACCGTAACCCATC 191
Qy 579 AAACATTGACTGGGCTACTGATGGGTAGAAATCTTCTTGAAGCAACTAGTTTCTTTGC 638
Db 192 TAACTCTCCCTGGGAGAACTAGGCATTGACCTTGATTCGAGGAACCGGAGTTTGT 251
Qy 639 TAAAAAGAAAGCTGCTGAAAAACACTTACATGCTAACCGTGTCTAAAAAGTTGTTATCAC 698
Db 252 GCACAGAGAAGGTGCTGGAAGACATTCAGCT---GGAGCCAAAGAGGTTCTTGAATAC 308
Qy 699 AGCTCTGCTGGAACGACGTTAAACAGAGTTGTTTCAACACTAACACGACATCTTTGA 758
Db 309 TGCACCTGTAAGAGAGATATCCCAACTTACGTTTGTGTCATGCTGACITTCAG 368
Qy 759 CGGTACTGAACAGATTATCTCAGGTCTTCATGTACTACAACTGTTTAGTCTCTATGCC 818

```



```

Db 369 CCATGAAGATCAATCATCAGAACGCTCTGTACTACTAACTGTCTCGTCCATCGT 428
Qy 819 TAAAGCTCTTACGATGATTTGTTATCCAAAAGGTTCTTATGACTACAATCCACGTTA 878
Db 429 CAAAGTCTTTCACAGAAATTCGGATCATCAAGGCAATGACAACTCACTCATA 488
Qy 879 TACTGGTACCAATATGATCTTACGAGGACCAACCGTGGTGGTACCTTCGTGCTCG 938
Db 489 CACTGGTACCAAGAGGCTGTGGATGCAAGCCAC-----GTGATCTAAGGAGAGCAAG 542
Qy 939 TGCTGGTGTGCAACATTTTCTTAACTCAACTGGTGTCTGCTAAAGCTATCGGTCTTGT 998
Db 543 AGCAGCTGTTTAAACATCGTTTCAACATCAACAGAGAGGCTAAAGCGGTGCTTGT 602
Qy 999 TATCCGAATGAAATGTAATCTGATGGTGTGCTGCAACAGTGTCTCTGTTTCAACTGG 1058
Db 603 CTTCCCTAACCTCAAGGAAAGCTCAAGGATCGCATTTGGGTGTCACAACTCCCAAGT 662
Qy 1059 ATCAGTAACTGAGTGGTGTGTAACCTTTGATAAAGCTTTTCTGTTGACGAAATCAAGC 1118
Db 663 CTCGGTGGTGTGATTTGTGCAAGTCTCCAGAGAGACTTTTGTGAGAGAGTCAAGC 722
Qy 1119 TGC---TATGAAGCTGTTTCAACAGCAGATTTTCGGTTTACACTGAAGATCCAAATTTTTC 1175
Db 723 TCGGTTACGAGGATGAGCTGCGAAGCTTAAAGGTATATCTCGATGTCTCGATGAGCC 782
Qy 1176 TTCAGATATCGAGGCTGTATAGCTTATCTTTTTCAGCGAATCAAACTAAAGTTAT 1235
Db 783 TCTTGTCTGTGCGACTTCAGGTGTCTGATGTGTCATCCACCATTTGTTCTCTCAC 842
Qy 1236 GGAAGTTACGGATCACAAATTTGGTTAAAGTTGTATCATGGTATGACAAATGCTTTTA 1295
Db 843 AATGTTATGGAGATGATATGTTTAAAGTATTGCTTGGTATGATATGATGATGAGGCTA 902
Qy 1296 CACTGCTCACTTGTTCGATACATTTGATGATTTTTCGAAATCGCTAAATTA 1347
Db 903 CTCTCAGAGTGTGATTTGGCTGACATTTGCTTCCCAATCACTGGAAGTGA 954

```

```

RESULT 3
AF075927
LOCUS
DEFINITION
  AF075927 Salmonella typhimurium LT2, Lambda DASH II genomic survey sequence.
ACCESSION
  AF075927
VERSION
  AF075927.1 GI:3320797
KEYWORDS
  GSS
SOURCE
  Salmonella typhimurium
  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
  Enterobacteriaceae; Salmonella.
REFERENCE
  1 (bases 1 to 944)
  Wong, R.M.-Y., Wong, K.K., Benson, N.R. and McClelland, M.
  Sample sequencing of a Salmonella typhimurium LT2 lambda library:
  comparison to the Escherichia coli K12 genome
  FEMS Microbiol. Lett. 173 (2), 411-423 (1999)
  99243757
  10227170
  Contact: McClelland M
  Molecular Biology
  Sidney Kimmel Cancer Center
  3099 Science Park Road, San Diego, CA 92121, USA
  Email: mclelland@lifesci.sdsu.edu
  Class: shotgun.
  Location/Qualifiers
    1..944
    /organism="Salmonella typhimurium"
    /mol_type="genomic DNA"
    /strain="LT2"
    /db_xref="taxon:602"
    /clone="B159-T7"
    /clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
FEATURES
  source

```

```

/note="Vector: Lambda DASH II; sequenced using Li-Cor
sequencer"
ORIGIN
Query Match 14.1%; Score 190.6; DB 28; Length 944;
Best Local Similarity 57.7%; Pred. No. 4.8e-38;
Matches 469; Conservative 0; Mismatches 314; Indels 30; Gaps 6;
Qy 497 TCGATGGTACTGTGAAGTTAAAGAGGTGGATTCGAAGTTAAACGTCATATTTGTTAAAG 556
Db 65 TCGACGGTACTGTGAAGTTAAAGAGGTGGATTCGAAGTTAAACGTCATATTTGTTAAAG 124
Qy 557 TTTCTGCTGAACCGCAACAGCAAAACATTAAGTGGCTACTGATGGCTAGAGGATAGTTC 616
Db 125 TTACCGCTGAACCGGATCCGGCTAACTGAAATGGG-ACGAATTTGTTGAGTAGTGG 183
Qy 617 TTGAAGCAACTAGTTCCTTTGCTAAAGAGAGCTGCTGAAAGCACTTACATGCTAAAG 676
Db 184 CTGAAGCTACCGGTATCTTCTGACTGAGGAAACCGGCTAAACACATCACCGCT---G 240
Qy 677 GTGCTAAAAAGTTTATCACAGCTCCTGGTGGAAACGAGCTTTAAACAGTTTGTTC 736
Db 241 GCGGAAAAAGTTTCTGACGGTCCGCTCTAAAGCAACACCCCAATTTGTTAAAG 300
Qy 737 ACCTAACACGACATTTTTCAGCGTACTGAAACAGTATCTCAGGTGCTTCACTACTA 796
Db 301 GCGTAACTTTGACAAATACGAAGGCGAGACATCGTT---TCCAAACGCTTCTCGACCA 357
Qy 797 CAACCTGTTTAGCTCTCTATGGCTAAAGCTCTTCAAGATGCATTTGGTATCCAAAAGTC 856
Db 358 CCAACTGCTTGGCGCGCTGGCTAAAGTTATCAACGACAACTTCGGCATCATCGAAGCTC 417
Qy 857 TTATGACTCAATCCACGCTTATCTAGTGTGACCAATGATCTTTCAGCGGACCAACCGTG 916
Db 418 TGAATGACTACTGTTTACGCGACTACCGCAACCCAGAAAAACCGTTGACGGCCGCTCTACA 477
Qy 917 GTGGTGACCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 976
Db 478 AA---GACTGGCGCGCGCGCGCGCGCATCTCAGAACATCATCCCGTCTCTACTGGCG 534
Qy 977 CTGCTAAAGCTATCGGTCTTGTATCCAGAAATGAATGTTAAACTTGTGTCGTCGTCAC 1036
Db 535 CTGCTAAAGCGGTAGGTAAAGTACTGCGGAACCTGAATGCAAACTGACTGTTATGCGCT 594
Qy 1037 AAGCTGTTCTTTCACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1096
Db 595 TCGCGGTTCTCTACTCCGAAAGCTATCCGTTGTTGACCTGACCCGTTCTGCTCGAAAAAGCGG 654
Qy 1097 TTTCTGTTGACGAAATCAACGCTGCTATGAAAGCTGCTTC-----AAACGACAG 1145
Db 655 CTACCTACGAGCAGATCAAGCTGCTGTTAAAGCTGTCGCGAAGNACATGAAGGCGT 714
Qy 1146 TTTGGTTACACTGAAGATCCAAATGTTTCTTTCAGATATCGTAGGGGTGTCATACGGTTC 1205
Db 715 GCTGGGTTACACCGAAGACGAGCTTGTATCTACCGATTTTCAACGGTGAAGTATGCACTTC 774
Qy 1206 ATTGTTTGGCGCAACTCAAACTAAAGTTATGGAAGTTGAGGATCACAATTTGTTAAAGT 1265
Db 775 CGTGTTCGATGCTAAAGCAGGCTTCGGCTGAACGACAA-----TTCGTGAAACT 825
Qy 1266 TGTATCATGCTATGACAAATGAAATGTTTACAC 1298
Db 826 GGCTCTCTGTCAGATAACGAAACCGGTTACTC 858

```

```

RESULT 4
CK261340
LOCUS
DEFINITION
  EST707418 potato abiotic stress cDNA library Solanum tuberosum cDNA
  clone POAB929 5' end, mRNA sequence.
ACCESSION
  CK261340
VERSION
  CK261340.1 GI:39818318
KEYWORDS
  EST.

```


SOURCE	Solanum tuberosum (potato)
ORGANISM	Solanum tuberosum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum.
REFERENCE	1 (bases 1 to 973)
AUTHORS	Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
TITLE	Generation of ESTs from abiotic stressed potato tissue
JOURNAL	Unpublished (2003)
COMMENT	Contact: Robin Buell The Institute for Genomic Research 9712 Medical Center Dr, Rockville, MD 20850, USA Email: potato-array@tigr.org Clones can be requested from TIGR via potato@tigr.org Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES	Location/Qualifiers 1. 973 /organism="Solanum tuberosum" /mol_type="mRNA" /cultivar="Kennebec" /db_xref="taxon:4113" /clone="FOAB929" /tissue_type="abiotic stress treated leaf and root tissue" /lab_host="DH10B-Tona" /clone_lib="potato abiotic stress cDNA library" /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1d, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."
ORIGIN	Query Match 14.0%; Score 189; DB 14; Length 973; Best Local Similarity 55.5%; Pred. No. 1.2e-37; Matches 432; Conservative 0; Mismatches 335; Indels 12; Gaps 3; QY 523 GGTGATTCGAAGTTAAGCGTCAATTTGTTAAAGTTTCTGCTGAACGGAACACGAGAAAC 592 Db 129 GATGGAAATCTGTGATGGAAAATATCAAGTCGTCTCAACCGTGACCTGTGAAC 188 QY 583 ATTGACTGGGCTACTGATGGCGTAGAAATGCGTTCTTTGAAGCAACTAGTTCTTTTGCTAAA 642 Db 189 CTCCATCGGGAGAACTTGGAATTGATTTGGTTCATAGAAGGTACCGGAGTGTTTGTAGAC 248 QY 643 AAAGAGCTGCTGAAAACACATTCATGCTAACGCTGTCTAAAAGTTGTTATCACAGCT 702 Db 249 AGAAGAGTTCGGGTAAACATCCAGGC---CGAGGCCAAGAGGTGCTCAACCGCC 305 QY 703 CCTGGTGGAAACGACGCTTAAACAGTTGTTTCAACACTAACACGACATTTCTGACGGT 762 Db 306 CCGCGAAAAGGTGATATCCCTACTTATGTTGTGTTGTTGTTCAATGCTGAACITTAACAACAT 365 QY 763 ACTGAAACAGTTATCTCAGTGTCTTCAGTACTACAACTGTTTAGCTCTCTATGGCTAAA 822 426 GTTCTTGACCAAAATTTGGAATTTATCAAGGGAACAATGACAACTACTCTCTTACACC 485 883 GGTGACCAAAATGATCCCTTGACGGACACACCGTGGTGGTACCTTCGTGCTCGTGTCT 942 486 GGTGACCAAAAGGCTTCCTTGATGCAAGCCACAGGGAT-----CTTAGACGTGCACGAGCT 539 943 GGTGCTGCAAAACATGTTCTTCACTCACTCACTGCTGCTCTAAAGCTATCGGTCTTGTATC 1002 540 GCAGCATTTGAACATAGTTCCAACTCACTGGTCTCTAAGGCTGTGGCTCTTTGTCTCTC 599 1003 CCAGAAATTTGAATGTTAACTTTGATGCTGCACAAAGTGTCTCTGTTGACGAATCAACGCTGATCA 1062 600 CCAAGCCTCAAGGGGAAACTCAATGGGATTCCTCCGTGTTCCACCCCTCAAGCTCTCA 659 1063 GTAACTGAGTTGGTTGTAATCTCTTGAATAAAACGTTTCTGTTGACGAATCAACGCTGCT 1122 660 GTTGTGGACCTTGTTCGTTCAAGTCACCAAGAAGACATTTGCTGGAAGTGAATGCTGCA 719 1123 ATGAAGCTGCTTCAAAACGACAGTTTCGGTTTACACTGAAGATCCAAATTTTCTTCAGAT 1182 720 TTCAGAGAAGCTGCTGATGAAGAACTCAATGGCATTCTAGCTGTCTGCGATGAACCACTT 779 1183 AT---CGTAGCGTGTTCATACCGTTCAATGTTTGACGCACTCAAACTAAAGTTATGGAA 1239 780 GTGTCACTTGACTTCCGGTGCAGTGTGTTCATCAACCACTTGAATCTTCACTCACCATG 839 1240 GTTGACGATCACAAATGGTTAAAGTTGTATCATGGTATGACAATGAATCTTTACAC 1298 840 GTCATGGGAGATGATCGTTAAAGTTATGCTTGGTATGACCATGAATGGGGTTACTC 898

using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The λ phage method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999)."

ORIGIN

Query Match 13.8%; Score 187.6; DB 28; Length 892;
 Best Local Similarity 54.3%; Pred. No. 2.8e-37; Indels 21; Gaps 4;
 Matches 476; Conservative 0; Mismatches 379;

447 TCAGAACAAATGACTGGGCACTACCGGTGTGAAATCGTTCTTGAATCGATGGTAC 506
 Db TCAGAACATACATGTTTACTTATTAATAATGATACCGGCCACAGCAATTCGAAGAAC 808

507 TGTGAGTTAAGAAGGTGGATTGCAAGTTAACGGTCAATTTGTTAAAGTTCTGCTGA 566
 Db TGTGAGCTGGAGAAATGCCATTATTGTTAATGGACACAAGATTGTTGCAAGCTGA 748

567 ACAGCAACAGCAAAATGACTGGGCTACTGATGGCGTAGAAATCGTTCTTGAAGCAAC 626
 Db AGAGATCCAGCTCAATTTGGATGGGAGCCCTTGGAGTTGATTGTTGATCAAC 688

627 TAGTTTCTTTGTAAGAAAGAGCTCTGAAACACACTTACATGCTTAACGGTGTGTA 686
 Db AGAGTTTCTCACTACTATTCCAAAGCTGAAGCTCATATTAAAG---GAGGAGCTAAGAA 631

687 AGTTGTTATCAGAGCTCTGGTGGAAACGACGTTAAACAGATTGTTTCAACACTAACCA 746
 Db AGTTATTTATTCAGCCCATCTG---CTGATGACCAATGTTTGTGTTGGAGTTAACAC 574

747 CGACATTTCTGCGSPACTGAACAGTTATCTCAGGTGCTTCATGATPACAAACGTGTT 806
 Db TGAGGCATACAAACCAAGAAATGAACATTATTTCAAATGTCATCATGACAAACCACTGCCT 514

807 AGCTCTATGGCTTAAGCTCTTCACATGATGATTTGGTATCCAAAGGCTTTATGACTAC 866
 Db TGACCACTGTCTAAAGTTATTAAACATGAATTTGGAATTTGTTGAAGGCTTTATGACAA 454

867 AATCCACGCTTATCTGCTGACCAATGATCTCTTGACGACCAACACCGTGTGGTCACT 926
 Db TATTCATGCTACTACTGCTACTCAAGACAGTTGATGACCATCAGGAAAGATTGGAG 394

927 TCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
 Db AGCTGTAGATGTGCAATGTGCT---AACATTATCCAGCTAGCAGCTGGAGCAGCTAAAGC 337

987 TATCGGTCTTGTATCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1046
 Db TGTGGAAGTTATTCCATCCCTTAATGGAAATCTTACTGTTATGCTCATTGAGATTGG 277

1047 TGTTCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1106
 Db AATCCAGATGTTTCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 217

1107 CGAAATCAAGCTGCTATGAAAGCTGCTTCA-----AAGCAGATTTGCTGTTA 1154
 Db ATCAATTAAGCTGCTATTAAAGAGATCAGAGAGAAATGAAGGAATTTCTTGGATA 157

1155 CACTGAAGATCCAAATTTCTTTTTCAGATATCGTAGCGGTGTCATACGGTTCATTTGTTGA 1214
 Db CACTGAAGACCTTGTGTTATCATCAGACTTTGTCATGATTAACAGATCATCAATTTTGA 97

1215 CGCACTCAAACTTAAAGTTATGAAGTTGACGATCAGATGTTGTTTAAAGTTGATCATG 1274
 Db TGAGGTGCTACTATTCTTTTGAACCCACACTTTGTTTAAATAGTTTTCATGTTGAGTAC 37

1275 GTATGACAAATGAAATGCTTTACACTGCTCAACTTGT 1310

36 ATCAAGAGGGAGACTTTTCCAGCACAGATTGT 1
 Db

BI433946 772 bp mRNA linear EST 30-APR-2003
 EST536707 P. infestans-challenged potato leaf, compatible reaction
 Solanum tuberosum cDNA clone PFCB36 5' sequence, mRNA sequence.
 BI433946
 VERSION BI433946.1 GI:15258636
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum.
 REFERENCE 1 (bases 1 to 772)
 AUTHORS Restrepo,S., Griffiths,H.M., Smart,C.D., Cho,J., Chiemiango,A.,
 Bougri,O., Buell,C.R., Rønning,C.M., Fry,W.E. and Baker,B.
 Generation of ESTs from Potato Leaves Challenged with Phytophthora
 infestans, Compatible Interaction
 Unpublished (2000)
 JOURNAL
 COMMENT Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics
 Institute. Orders can be made through the URL:
 http://genome.arizona.edu/orders/
 Seq primer: M13F-R.
 Location/Qualifiers
 1. 772
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="PFCB36"
 /tissue_type="leaf"
 /dev_stage="6 week old"
 /lab_hosts="SOLR"
 /clone_lib="P. infestans-challenged potato leaf,
 compatible reaction"
 /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
 XhoI; supplier: Cornell University, Fry lab; sequencing:
 The Institute for Genomic Research; Whole plants were
 challenged with 20,000 sporangia/ml of the compatible P.
 infestans isolate US 940480. Leaf tissue was collected at
 3, 6, 9, 12, 24, 48, 72 hours after inoculation and frozen
 in liquid nitrogen immediately upon removal. Kennebec
 plants showed first symptoms of infection at 48 hours
 after inoculation. NOTE: We cannot exclude the
 possibility that this sequence is actually derived from
 Phytophthora rather than potato."

ORIGIN

Query Match 13.8%; Score 185.4; DB 12; Length 772;
 Best Local Similarity 55.6%; Pred. No. 9.8e-37;
 Matches 423; Conservative 0; Mismatches 326; Indels 12; Gaps 3;

554 AAGTTTCTGCTGAACGGGACCAACCAACATTTGACTGGCTACTGATGGCGTAGAAATCG 613
 Db AAGTCGCTCCAAACCGTGACCCCTGTGAACCTCCCATGGGAGAACTTGGAAATGATTGG 67

614 TTCTTGAAGCAACTAGTATTTCTTCTTAAAGAAAGCTGCTGAAAAACACTTACATGCTA 673
 Db TCATAGAAGGTACCGGAGTGTGTTAGACAGAGAAGGTGCCGGTAAACACATCCAGGC-- 125

674 ACGETGTCTAAAAAGTTGTTTATACAGCTCTCTGCTGGAACAGCTTTAAAAACAGTTGTT 733
 Db -CGAGCCAGAGAGTGTCTCTATCCGCCGCCCGGAAAGGTGATATCCCTACTTATGTTG 184

734 TCAACACTAACCCAGCACATTCTTCCGCTACTGAAACAGTTATCTCAGGTGCTTCATGTA 793

QY 1183 AT---CGTAGGGGTGCATACGGTTCATTTGTTAGCGCACTCAAACTAAAGTTATGAA 1239
 Db 854 GTGTCAAGTTGACTTCGGGTGCGAGTGATGTCATCAACCAATGATCTTCCTCACTCCACATG 913
 QY 1240 GTTTCAGCGGATCAAAATGTTAAAGTTGATCATGGTATGAC 1281
 Db 914 GTCATGGGAGATGACATGGTTAAAGTATGCTTGGTATGAC 955

RESULT 8
 CK272330/c
 LOCUS
 DEFINITION
 EST718408 potato abiotic stress cDNA library Solanum tuberosum cDNA
 clone POAD453 3' end, mRNA sequence.
 CK272330
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Solanum tuberosum (potato)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 1 (bases 1 to 891)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum.
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
 Generation of ESTs from abiotic stressed potato tissue
 Unpublished (2003)
 Other ESTs: EST718407
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: GTA ATA CGA CTC ACT ATA GGG C.
 Location/Qualifiers
 1..891
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POAD453"
 /tissue_type="abiotic stress treated leaf and root tissue"
 /lab_host="PH108-Tona"
 /clone_lib="potato abiotic stress cDNA library"
 /note="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: Solanum tuberosum var. Kennebec plants were
 grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to
 four separate sets of plants. Set 1 involved saturation of
 the soil with 150 mM NaCl and tissues were harvested at
 following application of the salt stress (leaves: 2hr,
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
 Set 2 were grown under the standard conditions and then
 were water stressed by withdrawal of further watering
 applications. Drought stressed plants were harvested after
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
 and 5d). Set 3 were grown under the standard conditions
 and then were cold stressed by placement at 4 C. Cold
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d. Set 4 were grown under the standard conditions and
 then were heat stressed by placement at 35 C. Heat
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d and 4d and heat-stressed roots were harvested at 6 hr,
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and
 equal RNA from each tissue and stress was pooled to
 construct the cDNA library. RNA sample."

ORIGIN
 Query Match 13.3%; Score 179.2; DB 14; Length 891;
 Best Local Similarity 55.7%; Pred. No. 4, 1e-35;
 Matches 409; Conservative 0; Mismatches 313; Indels 12; Gaps 3;

QY 617 TTGAAGCAACTAGTTCTTTGCTAAAAAGAGCTGCTGAAAAACACTTACATGCTAACG 676
 Db 889 TAGAAGGTACCGGAGTGTGTTGTAGACAGAGAGGTGCGGTAAACACATCCAGCG---CG 833
 QY 677 GTGCTAAAAAGTTGTTATCACAGCTCTCTGTTGGAACAGAGCTTAAACAGTTGTTTCA 736
 Db 832 GAGCAAGAAGGTGCTCATCACCGCCCCCGGAAAGGTGATATCCCTACTTATGTTGTTG 773
 QY 737 ACACCTAACCCAGCATTTCTTGAGCGTACTGAAACAGATTATCTCAGGTGCTTCATGTA 796
 Db 772 GTGCTAATGTGAATTTTACAACCATGATGACCTATCATCAGCAATGCTCTTTCACCA 713
 QY 797 CAAACTGTTTAGTCTCTATGGCTAAAGCTCTTCCAGATGCAATTTGGTATCCAAAAAGT 856
 Db 712 CCAACTGCTTGTCTCTTCTGTCGAAGTCTTGTACCCAGAAATTTGGAATTTATCAAGG 653
 QY 857 TTATGACTACATCCAGCTTATCTGTTGACCAAAATGATCCTTGACGGACCCACCGCTG 916
 Db 652 CAATGAACAATCTACTCTTTACACCGGTGACCAAGGCTTCTTGTATGCAAGCCAC--- 597
 QY 917 GTGTGACCTTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
 Db 596 --AGGATCTTAGACGTGACAGAGCTGACGATTTGAACATAGTTTCCAACTCAACTGGT 539
 QY 977 CTGCTAAGCTATCGTCTTGTGTTATCCAGAAATGAAATGTAATGTAATGTAATGTAAT 1036
 Db 538 CTGCTAAGCTGCTGCTCTTGTCTCTCCAAAGCTCAAGGGGAACTCAATGGGAATGCC 479
 QY 1037 AACGTGTTCTGTTCCAACTGGATCACTGAGTTGGTGTAACTCTTTGATAAAAAACG 1096
 Db 478 TCCGTGTTCCCAACCCCTAAAGCTCTCACTGTTGACCTTGTCTCAAGTCACCAAGA 419
 QY 1097 TTTCTGTGACGAATCAACGCTGCTGATGAAGCTGCTTCAACGACAGTTTGGTTTACA 1156
 Db 418 CAITTTGCTGAGGAAGTGAATGCTGCTGATGAGAAAGCTGCTGATGAGAACTCAATGG 359
 QY 1157 CTGAAGATCCAAATTTGTTCTTTCAGATAT---CGTAGGCGTGTCTATACGGTTCA 1213
 Db 358 TTCTAGCTGTCTGCGATGAACCACTTGTGTGCTGAGCTTGTCTCCGTCAGTGTGTCAT 299
 QY 1214 ACGCACTCAAACTPAAAGTTATGGAAGTTGACGATCACAATGGTTAAAGTTGATTCAT 1273
 Db 298 CAACCAATGATTTCTTCACTCACCATGCTCATGGAGATGATGATGATGATGATGATG 239
 QY 1274 GGTATGCAATGAATCTTCTTACACTGCTCAACTGTTGTTTCGTACACTTGAATTTTCA 1333
 Db 238 GGTATGCAATGAATGAGTGGGTACTTCACAGAGGGTTGTTGATCTTGTGATGATTT 179
 QY 1334 AAATCGCTAAATAA 1347
 Db 178 ACCAGTGAATGA 165

RESULT 9
 CK272824
 LOCUS
 DEFINITION
 EST718902 potato abiotic stress cDNA library Solanum tuberosum cDNA
 clone POAD755 5' end, mRNA sequence.
 CK272824
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Solanum tuberosum (potato)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 1 (bases 1 to 973)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum.
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
 Generation of ESTs from abiotic stressed potato tissue
 Unpublished (2003)
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GNG ACA CTA TAG.
Location/Qualifiers

FEATURES

source
1. 973
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAD755"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="potato abiotic stress cDNA library"
/note="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1 d,
2d, and 4d) were grown under the standard conditions and then
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

ORIGIN

Query Match 13.2%; Score 178.2; DB 14; Length 973;
Best Local Similarity 58.3%; Pred. No. 7.7e-35;
Matches 353; Conservative 0; Mismatches 243; Indels 9; Gaps 2;
QY 523 GGTGATTTCGAAGTTAACGGTCAATTTGTTAAAGTTCTTGTGAACCGGAAACGCAAC 582
DB 378 GATGGAATCTCTGTCGATGGAAATATCATCAAGTCGTCTCAACCGTGACCCCTGTGAAC 437
QY 583 ATTGACTGGGCTACTGATGGGCTAGAAATCGTTCTTGAAGCACTAGTCTTTTGTCTAAA 642
DB 438 CTCCCATGGGAGAACCTTGGAGTTGATTTGTCTATAGAGGTACCGAGTGTTCGTAGAC 497
QY 643 AAAGAAGCTGTGAAAAACACTTACATGCTAACCGTGTCTAAAAAGTTGTTATCACAGCT 702
DB 496 AGAGAAGGTGCGGTAAACACATCCAGGC---CGGAGCAAGAGGTGCTCATCCCGCC 554
QY 703 CTTGTTGGAACGAGCTTAAACAGTTGTTTTCACACTTACCAGCATCTTTGACGGT 762
DB 555 CCGGAAAGGTGATATCCCTACTTATGTTGTGTGTCATGCTGAACCTTACACCAT 614
QY 763 ACTGAAACAGTTATCTCAGGTCCTTCACTGATCTACAAACTGTTTAGCTCTTATGGCTAAA 822
DB 615 GATGAACCTATCATCAGCAATGCTCTTGACCACTCACTGCTTGTCTCTTTCGTCAAG 674
QY 823 GCTCTTACGATGATTTGGTATCCAAAAGTCTTATGACTACATCCAGCTTATCT 882
DB 675 GTTCTTGACAGAAATTTGGAATATCAAAAGGAAACAATGACAACTACTCATCTTTACACC 734
QY 883 GTTGACCAATGATCTTTCAGCGGACCAACCGTGTGTGACCTTTCGTCGTCGTGCT 942
DB 735 GTTGACCAAGGCTTCTTGATCGAAGCCACAGGAT-----CTTAGAGCTGCACGAGCT 788
QY 943 GTGTGCTGAAAACATTGTTCTTAACCTCACTGCTGTCTTAAAGCTATCGGTCTTGTATC 1002
DB 789 GCAGCACTCAACATAGTTCCAACTCACTGCTGTCTTAAAGCTGTGGCTCTTGTCTC 848

QY 1003 CCAGAATTGAATGTAAACTTGTGCTGCTGCACAAAGTGTTCCTGTTCCAACTGGATCA 1062
DB 849 CCAAGCCTTAAGGGGAAACTCAGGCAATGCTCTCCGGTGTTCACCCCTCAAGCTCTCT 908
QY 1063 GTAACGTAGTTGGTTGTAACCTCTTGATAAAAAACGTTTCTGTTGACGAATCAACGCTGCT 1122
DB 909 GTTGTGACCTTGTGCTTTCAGTCACCAAGAGACATTTGCTGAGGAAGTGAATGCTGCA 968
QY 1123 ATGAA 1127
DB 969 TTCAA 973

RESULT 10

AI779267
LOCUS AI779267
DEFINITION EST260146 tomato susceptible, Cornell Lycopersicon esculentum cDNA clone cLES7N21, mRNA sequence.
ACCESSION AI779267
VERSION AI779267.1 GI:5277308
KEYWORDS EST
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 641)
D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,
Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y.,
Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,
Giovannoni, J.J. and Martin, G.B.
Generation of ESTs from pseudomonas susceptible tomato
Published (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.
Location/Qualifiers
1. 641
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Ril-13 (Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="cLES7N21"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato susceptible, Cornell"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLES - Tomato Pseudomonas Susceptible EST Library.
Directionally cloned cDNAs inserted into pBluescript
SK(-) at 5' end with EcoRI and 3' end with XhoI site"

FEATURES

source
1. 641
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Ril-13 (Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="cLES7N21"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato susceptible, Cornell"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLES - Tomato Pseudomonas Susceptible EST Library.
Directionally cloned cDNAs inserted into pBluescript
SK(-) at 5' end with EcoRI and 3' end with XhoI site"

ORIGIN

Query Match 13.2%; Score 177.8; DB 9; Length 641;
Best Local Similarity 58.1%; Pred. No. 8.4e-35;
Matches 355; Conservative 0; Mismatches 247; Indels 9; Gaps 2;
QY 523 GGTGATTTCGAAGTTAACGGTCAATTTGTTAAAGTTCTTGTGAACCGGAAACGCAAC 582
DB 23 GATGGCATCTCTGTCGATGAAAAGTCATCCAAAGTCGTCTCCAAACCGTGAC 82
QY 583 ATTGACTGGGCTACTGATGGGCTAGAAATCGTTCTTGAAGCACTAGTCTTTTGTCTAAA 642
DB 83 CTCCCATGGGAGAACCTTGGAGTTGACTTAGTCATAGAAGGTACCGAGTTTTGTAGAC 142
QY 643 AAAGAAGCTGCTGAAAACACTTACATGCTTAACGGTGTCTTAAAGTGTATTTATCACAGCT 702
DB 143 AGAGAAGGTGCGGTTAAACACATCCAGGC---CGGAGCAAGAGGTGCTCATCCCGCC 199

703 CTGTGGGAAACGACGTTAAACAGTGTGTTTCAACACTAAACGACAACTTTGACGGT 762
 200 CCGGAAAGGTGACATCCCTACTTATGTTGTGTTGCTCAATGCTCAACTTTACAGCCAT 259
 763 ACTGAAACAGTATCTCAGGTGCTTCATGCTACTACAACTGTTAGCTCCTATGCTAAA 822
 260 GATGAACCTATCATCAGCAATGCTCTTGTACCACTGCTGCTGCTCTTGTCTCAG 319
 823 GCTCTTCAAGATCAATTTGTTGTTTCAAAAGGTTTATGACTACATCAACGCTTATCT 882
 320 GTTCTTGACCAAAATTTGGAATTAATCAAGGAAACATGCACTACTCACTTTACACC 379
 883 GGTGACCAATGATCTTACGAGGACACACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 942
 380 GGTGACCAAGGCTTCTTGTATGCAAGCCACAGGGAT-----CTTAGAGCTGACAGCT 433
 943 GGTGCTGCAAAACATTTGCTTCACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1002
 434 CGAGCACTGACATGATTTCAACCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 493
 1003 CCAGAAATGAATGGTAACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062
 494 CCAAGCTCAAGGGGAACTCAACGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 553
 1063 GTAACCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1122
 554 GTTGTGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
 1123 ATGAAGCTGC 1133
 614 TTCAGAGGCG 624

RESULT 11
 BI422981
 LOCUS
 DEFINITION EST533647 tomato callus, TAMU Lycopersicon esculentum cDNA clone
 CLEF73C9 5' end, mRNA sequence.
 ACCESSION BI422981.1 GI:15197841
 VERSION
 KEYWORDS
 SOURCE Lycopersicon esculentum (tomato)
 ORGANISM
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 783)
 Alcala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T.,
 Holt, I.E., Liang, F., Upton, J., Craven, M.B., Bowman, C.L., Ahn, S.,
 Roming, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and
 Giovannoni, J.
 Generation of ESTs from tomato callus tissue
 Unpublished (1999)
 Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>.
 Location/Qualifiers
 1..783
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA96"
 /db_xref="taxon:4081"
 /clone="CLEF73C9"
 /tissue_type="callus"
 /dev_stage="25-40 days old"
 /lab_host="XLI-Blue MRF"
 /clone_lib="tomato callus, TAMU"
 /notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
 XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
 of seedlings 7-10 days post-germination were excised, cut
 at both ends and placed on MS medium with no selection.

ORIGIN
 Query Match 13.2%; Score 177.8; DB 12; Length 783;
 Best Local Similarity 58.1%; Pred. No. 9e-35;
 Matches 355; Conservative 0; Mismatches 247; Indels 9; Gaps 2;
 523 GGTGATTCGAAGTAAACGGTCAATTTGTTAAAGTTTCTGCTGACGCGAACGACAAAC 582
 73 GATGCACTCTCTGTCGATGGAAGAGTCAATCCAGTCTGCTCTCAACCGCTGACCTGTGAAC 132
 583 ATTGACTGGGCTACTGATGGGTAGAAATCGTTCTTGAAGCAACTAGTTCTTTTCTGCTAAA 642
 133 CTCCATGCGGAGAACCTTTGGAGTTGACTTAGTCAATAGAGGTACCGAGTTTGTGTAGAC 192
 643 AAAGAAGCTGCTGAAAAACACTTACATGCTAAACGGTCTAAACAAAGTTTGTATCACAGCT 702
 193 AGAGAAGGTGCGCGGTAAACACATCCAGC---CGAGCCAAAGAGGTGCTCATCACGCC 249
 703 CTGCTGGAACAGCGTTTAAACAGTTGTTTCAACACTACACGACATTTCTTGACGGT 762
 250 CCGGAAAGGTGACATCCCTTACTTATGTTGTTGGTCAATGCTGAACCTTTACAGCAT 309
 763 ACTGAAACAGTTATCTCAGGTGCTTCACTACTCAAACTGTTTGTAGCTCTTATGGCTAAA 822
 310 GATGAACCTATCATCAGCAATGCTCTTGTACCAACCAACTGCTTGTCTTTCGTCAG 369
 823 GCTCTTACAGTGCATTTGTTATCAAAAAGTCTTATGACTACAACTCCAGCTTATCT 882
 370 GTTCTTGACCAAAATTTGGAATTTATCAAGGGAACAATGACACACTACTCACTCTTTACACC 429
 883 GGTGACCAATGATCTTACGCGACCAACCGTGTGCTGACCTTCTGCTGCTGCTGCTGCT 942
 430 GGTGACCAAGGCTTCTGATGCAAGCCACAGGGAT-----CTTAGAGCTGACAGCT 483
 943 GGTGCTGCAACATGTTCTTCACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1002
 484 GCAGCACTGAACATAGTTTCAACCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
 1003 CCAGAAATGATGTTGTTAACTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062
 544 CCAAGCTCAAGGGAACACTCAACGGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603
 1063 GTAACCTGAGTGTGTTGTTAACTTCTTGAATAAAACGTTTCTGTTGACGAAATCAACGCTGCT 1122
 604 GTTGTGGACCTTGTGTCGAAGTCAACGAAGAAGACATTTGCTGAGGAAGTGAATGCTGCA 663
 1123 ATGAAGCTGC 1133
 664 TTCAGAGGCG 674

RESULT 12
 AW267163
 LOCUS
 DEFINITION YAS1 gastric carcinoma cell GC7901 Homo sapiens cDNA, mRNA
 sequence.
 ACCESSION AW267163
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 530)
 Zhao, J.R., Yan, X.J., Han, F.C., Cui, D.X., Hou, Y., Yan, Q.J. and
 Su, C.Z.
 Gastric associated differentially expressed gene mRNA sequence
 Unpublished (1999)
 Contact: Zhao JR
 Institute of Genetic Diagnosis, the Fourth Military Medical
 University

AUTHORS
 D' Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,
 Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y.,
 Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,
 Giovannoni, J.J. and Martin, G.B.
TITLE
 Generation of ESTs from *Pseudomonas* susceptible tomato
JOURNAL
 Unpublished (1999)
COMMENT
 Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source
 1..626
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="R11-13 (Rio Grande x Money Maker)"
 /db_xref="taxon:4081"
 /clone="cLES20113"
 /tissue_type="leaf"
 /dev_stage="4-week old"
 /lab_host="SOLR"
 /clone_lib="tomato susceptible, Cornell"
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
 XhoI; cUES - Tomato *Pseudomonas* Susceptible EST library.
 Directionally cloned cDNAs inserted into pBluescript.
 SK(-) at 5' end with EcoRI and 3' end with XhoI site"

ORIGIN
 Query Match 13.0%; Score 175.4; DB 9; Length 626;
 Best Local Similarity 58.3%; Pred. No. 3.5e-34;
 Matches 349; Conservative 0; Mismatches 241; Indels 9; Gaps 2;

 QY 535 GTTAACGGTCAATTTGTTAAAGTTCTCTGTGACGCGAACCGAACATGTGACTGGGCT 594
 Db 1 GTCGATGGAAAGTCATCAAGTCGCTCTCAACCGTGACCCCTGTGAACCTCCCATGGGA 60

 QY 595 ACTGATGGGTAGAAATCTCTTGAAGCAACAGTATTTCTTCTCTAAAAAGAAAGCTGCT 654
 Db 61 GAACTTGGAGTTGACTTAGTCATAGAAAGGTACCGGAGTTTTGTAGACAGAGAGTGCC 120

 QY 655 GAAAAACACTTACATGCTAACCGTGTCTAAAAAGTTGTTATCAAGCTCTCTGTGGAAAC 714
 Db 121 GGTTAAACATCCAGC--CGGAGCCAAAGAGTGTCTATCACCGCCCCGGAAAGGT 177

 QY 715 GAGCTTAAACAGTTGTTTTCACACTTAACCCAGCATCTTTCAGCGTACTGGAACAGTT 774
 Db 178 GACATCCCTACTTATGTTGTGTGTCATGCTGGAATTTACAGCCATGATGAACCTATC 237

 QY 775 ATCTCAGGTGCTTCATGTACTACAACTGTTTAGCTCCTTAGCTAAAGCTCTTCACGAT 834
 Db 238 ATCAGCAATGCCTCTGTGTACCAACCACTGCCTTCTCTTCGTCGTCGTCGTCGTCGTC 297

 QY 835 GCATTTGGTATCCAAAAGGCTTTATGATACATCCACGCTTATATCTGTGACCAATG 894
 Db 298 AAATTTGGAATTATCAAGGGAACAATGACAACTACTCACTCTTACCGGTGACCAAGG 357

 QY 895 ATCTTTGACGGACCAACCGTGGTGGTGAACCTTCGTGCTCGTGTGCTGTCGCAAC 954
 Db 358 CTTCTTGATGCAAGCCACAGGAT-----CTTAGAGTGCACGAGTGCAGCACTGAAC 411

 QY 955 ATTGTTCTCTAACTCAACTGCTGCTGTCTAAAGCTATCGGTCTTGTTATCCAGAAATTGAAT 1014
 Db 412 ATAGTTCACACCTCAACTGCTGTCTTAAGGCTGTGGCTCTGTGTTCTTCCCAAGCCTCAAG 471

 QY 1015 GGTTAACTTGATGGTGTGCAACACGTTGCTCTTCCCTGTTCCCACTGATCAGTAACTGAGTTG 1074
 Db 472 GGGAACTCAACGGCAATTGCCCTCCGTTGTTCCACCCCTTAACGCTCGTTGTGGACCTT 531

 QY 1075 GTTGTAATCTTGATAAAAAAGCTTTCTGTTGACGAAATCAACGCTGTGTGATAAGAGCTGC 1133
 Db 532 GTTGTGCAAGTCCCAAGAGACATTTGCTGAGGAAGTGAATGCTGTGATTCAGAGAGGC 590


```

RESULT 14
BQ045826
LOCUS
DEFINITION
    BQ045826 771 bp mRNA linear EST 10-MAR-2003
    EST594944 P. infestans-challenged potato leaf, incompatible
    reaction Solanum tuberosum cDNA clone BPL113A7 5' end, mRNA
    sequence.
ACCESSION
    BQ045826
VERSION
    BQ045826.1 GI:19819812
KEYWORDS
    EST.
SOURCE
    Solanum tuberosum (potato)
ORGANISM
    Solanum tuberosum
REFERENCE
    Zhang, P., Hernandez, M., Tornqvist, C.-E., Witz, U., Loukolanov, A.,
    Rangel, P., Haberlach, G.T., Karamycheva, S.A., Tsai, J., Chiemiango, A.,
    Bougri, O., Buell, C.R., Ronning, C.M., Helgeson, J. and Baker, B.
    Generation of ESTs from Potato Leaves Challenged with Phytophthora
    infestans, incompatible interaction (2002)
    Unpublished (2002)
    Contact: Robin Buell
    The Institute for Genomic Research
    9712 Medical Center Dr. Rockville, MD 20850, USA
    Email: potato-array@igr.org
    This clone can be obtained from the University of Arizona Genomics
    Institute. Orders can be made through URL:
    http://genome.arizona.edu/orders/
    Seq primer: T3.
FEATURES
    source
        1..771
        /organism="Solanum tuberosum"
        /mol_type="mRNA"
        /cultivar="Kennebec"
        /db_xref="taxon:4113"
        /clone="BPL113A7"
        /tissue type="leaf"
        /dev stage="6 week old"
        /lab_host="SOLR"
        /clone_lib="P. infestans-challenged potato leaf,
        incompatible reaction"
        /notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
        XhoI; supplier: UC Berkeley, PGEC; sequencing: The
        Institute for Genomic Research. Whole plants were
        challenged with 450,000 sporangia/ml P. infestans isolate
        US-1 (US940501) in Biotron (Madison, Wisconsin). Leaf
        tissue was collected at 1, 2, 5, 12, and 24 hours
        post-challenge and frozen in liquid nitrogen immediately
        upon removal. Kennebec plants showed no signs of HR.
        Katahdin plants (susceptible to P. infestans US-1) were
        used as controls and showed infection. NOTE: We cannot
        exclude the possibility that this sequence is actually
        derived from Phytophthora rather than potato."
ORIGIN
    Query Match 12.9%; Score 174; DB 12; Length 771;
    Best Local Similarity 55.3%; Pred. No. 8.6e-34;
    Matches 426; Conservative 0; Mismatches 330; Indels 14; Gaps 4;
    QY 523 GGTGGATTTCGAGTTACGTTCAATTTGTTAAAGTTTCTGTAACGCGAACCAGCAAC 582
    DB 6 GATGGAATCTCTGCGATGGAATAATCATCAAGTCGTCTCCACCGTGACCTGTGAAC 65
    QY 583 ATTGACTGGGCTACTGATGGCGTAGAAATCGTTCTTGAAGCAACTAGTTTCTTTGTAA 642
    DB 66 CTCCCATGGGAGAACTGGAGTTGATTGGTCATAGAGGTACCGAGTGTGTTGTAGAC 125
    QY 643 AAGAACTGCTGAAACACTTACATGCTAAAGGCTCTAAAGAGTTGTATCACAGCT 702
    DB 126 AGAAGAGTCCGGTAAACATCCAGGC--CGAGCCAGAGGGTCTCATCACGCC 182
    QY 703 CCTGGTGAACGACGCTTAAACAGTTGTTTCAACACTAACCCAGCATTTCTGCGGT 762

```

```

Db 183 CCCGAAAAGGTGATATCCCTACTTATGTTGTTGGTGCATGCTGAATTTACAACCAT 242
QY 763 ACTGAAAACAGTTATCTCAGGTGCTTCATGACTACAAAACCTGTTAGCTCTCTATGGCTAAA 822
Db 243 GATGAACCTATCATCAGCAATGCTCTTTGACCAACCAACTGCTTGTCTCTTCTGCTCAAG 302
QY 823 GCTCTTCAAGATGCAATTTGGTATCCAAAAGGTCTTATGACTCAAACTCCACGCTTATACT 882
Db 303 GTTCTTGACCAAGAAATTTGGAATTTATCAAGGAACAATGACAATCTACTCTTACACC 362
QY 883 GGTGACCAATGATCCTTTGACGGACCAACCGGTGGTGGTGAACCTTCGTCGTCGTGCT 942
Db 363 GGTGACCAAGGCTTCTTTGATGCAAGCCACA-----GGGATCTTAGACGTGCAAGAGCT 416
QY 943 GGTGCTGAAAACATGTTCTTAACCTCAACTGCTGCTGCTGAAGCTATCGGCTCTGTTATC 1002
Db 417 GCAGCACTCAACATAGTTCCAACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 476
QY 1003 CCAGAAATTTGAATGTTAAACTTTGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1062
Db 477 CCAAGCCTTAAAGGAAACTCAACGGCATTTGCTCCGTGTTCCACCCCTTAACGTCTCT 536
QY 1063 GTAACGTGTTGGTGTGTAACCTCTTTGATAA-AAACGTTTCTGTTGACGAAATCAACGCTGC 1121
Db 537 GTTGTGACCTTGTGCTTCAAGTCAACCAAGAACATTTTGTCTGAGGAAGTGAATGCTGTC 596
QY 1122 TATGAAGCTGCTCTCAACAGCAGCTTTGGTGTGCTGCAACAGCTGTTCTGTTTCCAACTGGATCA 1181
Db 597 ATTGAGAGGCTGCTGATAGGAACCTCAATGGCATTTCTATCTCTCTGATGAACCACT 656
QY 1182 TATCGTAGGCGTGT-----CATACGGTTCAATTTGTTGACGCAACTCAAACTAAAGTTATGG 1237
Db 657 CGTGTCAGTTGACTTCCGGTGCAGTGACGTGTCAATCAACCATGATTTCTTCACTCACC 716
QY 1238 AAGTTGACGGATCACAATTTGTTAAAGTTGATCATGCTATGACAATGAA 1287
Db 717 TGGTCATGGGAGATGACATGTTTAAAGTTATTGCTTGTGTAAGCAATGAA 766

```

```

RESULT 15
BQ045826
LOCUS
DEFINITION
    BQ045826 698 bp mRNA linear EST 12-DEC-2002
    MCS038C08 160608 Ice plant Lambda Uni-Zap XR expression library, 5
    days 0.5 M NaCl treatment, Crassulacean acid metabolism, phase 1 (2
    AM). Mesembryanthemum crystallinum cDNA clone MCS038C08 5', mRNA
    sequence.
ACCESSION
    BQ045826
VERSION
    BQ045826.1 GI:26563315
KEYWORDS
    EST.
SOURCE
    Mesembryanthemum crystallinum (common iceplant)
ORGANISM
    Mesembryanthemum crystallinum
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    Caryophyllales; Aizoaceae; Mesembryanthemum.
REFERENCE
    1 (bases 1 to 698)
    Cushman, J.C.
    An expressed sequence tag database for the common ice plant,
    Mesembryanthemum crystallinum
    Unpublished (1997)
    Contact: Cushman JC
    Department of Biochemistry
    University of Nevada
    MS200, Reno, NV 89557-0014, USA
    Tel: 775-784-1918
    Fax: 775-784-1650
    Email: jcushman@unr.edu
    PCR Primers
    FORWARD: T3 20mer
    BACKWARD: T7 21mer
    Plate: 038 row: C column: 08
    Seq primer: T3 20mer
    High quality sequence stop: 698.

```



```

FEATURES
  source      Location/Qualifiers
              1..698
               /organism="Mesembryanthemum crystallinum"
               /mol_type="mRNA"
               /db_xref="taxon:3544"
               /clone="MCS038C08"
               /tissue_type="leaf"
               /dev_stage="five-week-old"
               /clone_lib="ice plant lambda Uni-Zap XR expression
               library, 5 days 0.5 M NaCl treatment, Crassulacean acid
               metabolism, phase I (2 AM)."
               /note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site 1:
               EcoRI; Site 2: XhoI; Library construction was performed
               according to Stratagene's recommended protocol for the
               Lambda UnizapXR vector and cDNA synthesis kit."

ORIGIN
  Query Match      12.8%; Score 172.8; DB 14; Length 698;
  Best Local Similarity 57.0%; Pred. No. 1.7e-33;
  Matches 359; Conservative 0; Mismatches 262; Indels 9; Gaps 2;

Qy 514 GTTAAAGAGTGGATTGCGAAGTAAAGGTCATTTGTTAAAGTTTCTGCTGAACGGAA 573
Db   |||
Qy 50  GTTGGGATGATGCATCTCAGTTGATGGCAAAATCATCAAGTTGTCTCAGACAGAAC 109
Db   |||
Qy 574 CCAGCAACATTGACTGGCTACTGATGGCGTAGAATCGTTCTTGAAGCAACTAGTTTC 633
Db   |||
Qy 110 CCTGTCAACCTTCCTTGGAGAAATGGGCATCGATTGGTGATAGAGGACCGGAGTG 169
Db   |||
Qy 634 TTTGCTAAAAAGAGTGTCTGAAAAACACTTACATGCTAACGGTGTCTAAAAAGTTGTT 693
Db   |||
Qy 170 TTTGTGGACAGAGCGGTGCAGGTAAAGCACITGCAAGC---CGAGCTAAGAGGTCTC 226
Db   |||
Qy 694 ATCAGAGTCTCTGGTGGAAACGAGCTTAAACAGTTGTTTCAACACTAACCCGACATT 753
Db   |||
Qy 227 ATAAGTCTCAGGAAAGGGTGATATCCAACTATGTTGGTGGTGTCATGAGGAGAG 286
Db   |||
Qy 754 CTTGACGGTACTGAAACAGTTATCTCAGGTGCTTCATGCTACTACAAACTGTTTAGCTCCT 813
Db   |||
Qy 287 TATGACCCAGTGGAGACCATCATCAGTAACGCTCTCCGACCACTAACTGCTTGGCACCC 346
Db   |||
Qy 814 ATGGCTAAAGCTTTTCAGATGCAATTTGGTATCCAAAGAGTCTTTGACTACATCCAC 873
Db   |||
Qy 347 TTTGTCAAGGTCCTCGACCCAGAAATTTGGCATCATCAAGGGAACCATGACCACTACCCAC 406
Db   |||
Qy 874 GCTTATCTGTGGTACCAATGATCCTTCAGCGACCCACACCGTGGTGGTGCCTTCGTCGT 933
Db   |||
Qy 407 TCCTACACTGGTGACAGAGCTTCTAGCCAGTACCC-----GTGACCTAAGGCGT 460
Db   |||
Qy 934 GCTCGTGTGGTGTGCAAAACATTTCTTAACTCAACTGGTGTGCTGCTTAAAGCTATCGGT 993
Db   |||
Qy 461 GCAAGAGCTGCAGCCCTCAACATTGTCCCACTCACTGGTGCAGCCAGGCAAGTGGCT 520
Db   |||
Qy 994 CTTGTATCCAGAAATTCAGTGAATTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1053
Db   |||
Qy 521 TTGGTTCTGCCAGGCTCAAGGGCAAGCTCAACGGGATCGCTCTTCGGTGTCCGACCCCC 580
Db   |||
Qy 1054 ACTGATCAGTAACAGTGGTGGTGTGTTTAACTCTTGAATAAAAGCTTTCTGTTGACGAAATC 1113
Db   |||
Qy 581 AATGTCTCAGTGGTGTGACCTTGTCTCCAGGTCCACCAAGAGACCTTCGCTGAGGAGTG 640
Db   |||
Qy 1114 AACGCTGCTATGAAGCTGCTTCAAAAGAC 1143
Db   |||
Qy 641 AATGCTGCTTCAGAGAGGCTGCTGACAAAC 670
Db   |||

```

Search completed: March 31, 2004, 01:55:47
Job time : 3354 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 15:57:18 ; Search time 96.4776 Seconds
(without alignments)
1312.028 Million cell updates/sec

Title: US-10-650-369-22

Perfect score: 2278

Sequence: 1 MKKITGIILLAVILSAC.....EMSYAQLVRLTFYAKIAK 448

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2278	100.0	448	5	AAM50664 Streptoco
2	1656.5	72.7	336	5	AAM50665 Streptoco
3	1656.5	72.7	336	5	AAM50639 Streptoco
4	1655.5	72.7	336	5	ABP29960 Streptoco
5	1655.5	72.7	336	6	ABU46455 Protein e
6	1655.5	72.7	345	5	ABP29106 Streptoco
7	1652.5	72.5	336	2	AAR56486 Plasmin r
8	1652.5	72.5	336	4	AAR56681 Streptoco
9	1564.5	68.7	336	5	ABP30758 Streptoco
10	1559.5	68.5	336	5	AAM50667 Streptoco
11	1559.5	68.5	336	5	AAM50641 Streptoco
12	1557.5	68.4	336	5	AAM50666 Streptoco
13	1557.5	68.4	336	5	AAM50640 Streptoco
14	1535	67.4	335	6	ABU02516 S. pneumo
15	1535	67.4	339	6	ABU37576 Streptoco
16	1535	67.4	359	6	ABU46262 Protein e
17	1534.5	67.4	336	5	AAM50669 Streptoco
18	1534.5	67.4	336	5	AAM50643 Streptoco
19	1534	67.3	359	4	AAU38000 Streptoco
20	1530.5	67.2	336	5	AAM50668 Streptoco
21	1530.5	67.2	336	5	AAM50642 Streptoco
22	1521	66.8	337	6	ABU44330 Protein e
23	1491	65.5	333	2	ABW5089 Streptoco
24	1491	65.5	333	5	ABP54583 S. pneumo
25	1491	65.5	333	7	ADC45135 S. pneumo

26	1357.5	59.6	336	5	ABW55601 Lactococc
27	1282	56.3	333	4	AAU35254 Enterococ
28	1282	56.3	333	5	ABU29370 Protein e
29	1275	56.0	336	5	ABW53868 Lactococc
30	1270	55.8	333	7	ADC95142 E. faeciu
31	1270	55.8	333	6	ABU25342 Protein e
32	1224.5	52.9	336	5	ABW48810 Listeria
33	1205.5	52.9	336	6	ABU32568 Protein e
34	1203	52.8	334	6	ABU23595 Protein e
35	1166	51.2	335	6	ABU43810 Protein e
36	1145	50.3	334	6	ABU37737 Protein e
37	1141	50.1	357	6	ABP81005 N. gonorr
38	1141	50.1	357	6	ABU37077 Protein e
39	1138.5	50.0	336	4	AG83058 S. epider
40	1138.5	50.0	336	6	ABU42966 Protein e
41	1138.5	50.0	348	5	ABP40668 Staphyloc
42	1132.5	49.7	336	4	AAU36947 Staphyloc
43	1132.5	49.7	336	4	AAU34122 Staphyloc
44	1132.5	49.7	336	6	ABJ19195 Pathogen
45	1132.5	49.7	336	6	ABU16073 Protein e

ALIGNMENTS

RESULT 1
AAM50664
ID AAM50664 standard; protein; 448 AA.
XX
AC AAM50664;
XX
DT 29-AUG-2003 (revised)
DT 08-APR-2002 (first entry)
XX
DE Streptococcus GapC multiepitope fusion PolyGap4.
XX
KW PolyGap4; GapC; plasmin binding protein; epitope; infection; vaccine;
KW immunisation; mastitis; therapy.
XX
OS Streptococcus dysgalactiae.
OS Streptococcus agalactiae.
OS Streptococcus parauberis.
OS Chimeric.
XX
PN WO200196379-A2.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-CA000836.
XX
PR 12-JUN-2000; 2000US-0211247P.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Potter AA, Perez-Casal J, Fontaine M;
XX
FI WPI; 2002-098051/13.
XX
DR N-PSDB; ABA91327.
XX
PT Novel GapC multiple epitope fusion polypeptide comprising antigenic
PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
PT parauberis, or S. iniae GapC protein useful for treating mastitis in
vertebrates.
XX
PS Claim 9; Fig 6A-C; 116pp; English.
XX
CC The present sequence is that of a novel multiple epitope fusion protein,
CC designated PolyGap4, comprising the entire amino acid sequence of the
CC Streptococcus dysgalactiae GapC plasmin binding protein in addition to
CC unique amino acid sequences from the Streptococcus parauberis and
CC Streptococcus agalactiae GapC proteins. The multiple epitope protein is
CC produced in host cells transformed with an expression vector comprising a
CC chimeric gene (see ABA91327) encoding the protein. PolyGal4 is an example

CC of novel GapC multiple epitope fusion proteins of the invention that
 CC comprise epitopes from 1 or more of *S. dysgalactiae*, *S. agalactiae*, *S.*
 CC *parvaeris*, *Streptococcus uberis* and *Streptococcus iniae* (see AAM50665-
 CC 69). The multiple epitope fusion proteins are used in claimed vaccines
 CC for treating or preventing a bacterial infection in a vertebrate,
 CC especially a streptococcal infection, and particularly mastitis. They are
 CC also used in claimed methods of detecting *Streptococcus* antibodies. The
 CC multiple epitope protein is capable of eliciting broad immunity against a
 CC variety of streptococcal infections while minimising the number of
 CC antigens present in the final formulation and concomitantly reducing
 CC production costs. (Updated on 29-AUG-2003 to standardise OS field)
 XX
 SQ Sequence 448 AA;

Query Match 100.0%; Score 2278; DB 5; Length 448;
 Best Local Similarity 100.0%; Pred. No. 1.8e-180;
 Matches 448; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKIGIILLAVIILSACQANYGSMVKVINGRGRIGRLAFRIQNVGVEVTRIN 60
 DB 1 MKKIGIILLAVIILSACQANYGSMVKVINGRGRIGRLAFRIQNVGVEVTRIN 60
 QY 61 DLTPNMLAHLKYDTTQGRDGTVEVKEGFEVNGNFIKVSARDPENIDWATDGVIEI 120
 DB 61 DLTPNMLAHLKYDTTQGRDGTVEVKEGFEVNGNFIKVSARDPENIDWATDGVIEI 120
 QY 121 LEALGTVVKGDFGVNGKFIKVSARDPENIDWATDGVIEI 180
 DB 121 LEALGTVVKGDFGVNGKFIKVSARDPENIDWATDGVIEI 180
 QY 181 GQFVKVSAREPANIWDATDGVIEI 240
 DB 181 GQFVKVSAREPANIWDATDGVIEI 240
 QY 241 KTVVFNTHDILDTGTTVISGCTTNCLAPMAKALHDAFGIQKGLMTTHAYTGDMIL 300
 DB 241 KTVVFNTHDILDTGTTVISGCTTNCLAPMAKALHDAFGIQKGLMTTHAYTGDMIL 300
 QY 301 DGPHRGDLRRARAGANIVPNSGAAGAIGLVIPELNGKLDGAAQRPVPTGVTSLVV 360
 DB 301 DGPHRGDLRRARAGANIVPNSGAAGAIGLVIPELNGKLDGAAQRPVPTGVTSLVV 360
 QY 361 TLDKNVSVDEINAAKASNDSPGYTEDPIVSSDIVGVSGSLFDATQTKVMEVDGSQLV 420
 DB 361 TLDKNVSVDEINAAKASNDSPGYTEDPIVSSDIVGVSGSLFDATQTKVMEVDGSQLV 420
 QY 421 KVSVDNEMSYTAQLVRLTLEYFAKIAK 448
 DB 421 KVSVDNEMSYTAQLVRLTLEYFAKIAK 448

RESULT 2
 AAM50665
 ID AAM50665 standard; protein; 336 AA.
 XX
 AC AAM50665;
 XX
 DT 08-APR-2002 (first entry)
 XX
 DE Streptococcus dysgalactiae gapC plasmin binding protein DysGapC.
 XX
 KW DysGapC; GapC; plasmin binding protein; epitope; infection; vaccine;
 XX immunisation; mastitis; therapy.
 XX
 OS Streptococcus dysgalactiae.
 XX
 FN WO200196379-A2.
 XX
 PD 20-DEC-2001.
 XX
 PF 11-JUN-2001; 2001WO-CA000836.
 XX
 PR 12-JUN-2000; 2000US-0211247P.

XX (UYSA-) UNIV SASKATCHEWAN.
 PA Potter AA, Perez-Casal J, Fontaine M;
 XX
 PI WPI; 2002-098051/13.
 XX N-PSDB; ABA91328.
 DR
 DR Novel GapC multiple epitope fusion polypeptide comprising antigenic
 PT determinant of *Streptococcus dysgalactiae*, *S. agalactiae*, *S. uberis*, *S.*
 PT *parvaeris*, or *S. iniae* GapC protein useful for treating mastitis in
 PT vertebrates.
 XX
 PS Claim 8; Fig 1A-B; 116pp; English.

The present sequence is that of the GapC plasmin binding protein,
 CC DysGapC, of *Streptococcus dysgalactiae* ATCC 43078, an isolate from a case
 CC of bovine mastitis. The invention relates to novel GapC multiple epitope
 CC fusion proteins that comprise epitopes from 1 or more of *Streptococcus*
 CC *dysgalactiae*, *Streptococcus agalactiae*, *Streptococcus parvaeris*,
 CC *Streptococcus uberis* and *Streptococcus iniae* (see AAM50665-69). A claimed
 CC example is PolyGap4 (see AAM50664). Expression vectors and host cells for
 CC production of the multiple epitope fusion proteins are provided. The
 CC multiple epitope proteins are used in claimed vaccines for treating or
 CC preventing a bacterial infection in a vertebrate, especially a
 CC streptococcal infection, and particularly mastitis. They are also used in
 CC claimed methods of detecting *Streptococcus* antibodies. The multiple
 CC epitope proteins are capable of eliciting broad immunity against a
 CC variety of streptococcal infections while minimising the number of
 CC antigens present in the final formulation and concomitantly reducing
 CC production costs
 XX
 SQ Sequence 336 AA;

Query Match 72.7%; Score 1656.5; DB 5; Length 336;
 Best Local Similarity 79.6%; Pred. No. 4.9e-123;
 Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVKVGINGFGFRIGRLAFRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRDGTVEV 87
 DB 1 MVKVGINGFGFRIGRLAFRIQNVGVEVTRINDLTPNMLAHLKYDTTQGRDGTVEV 60
 QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEI 147
 DB 61 KEGGFEVNGNFIKVSARDPE 81
 QY 148 KDPEQIDWATDGVIEI 207
 DB 82 95
 QY 208 ATFFAKKEAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDTGTTVISGASCTTN 267
 DB 96 ATFFAKKEAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDTGTTVISGASCTTN 155
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGANIVPNSGA 327
 DB 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGANIVPNSGA 215
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDEINAAKASNDSPGYTE 387
 DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDEINAAKASNDSPGYTE 275
 QY 388 DFISSDIVGVSGSLFDATQTKVMEVDGSQLVSKVDNEMSYTAQLVRLTLEYFAKIA 447
 DB 276 DFISSDIVGVSGSLFDATQTKVMEVDGSQLVSKVDNEMSYTAQLVRLTLEYFAKIA 335
 QY 448 K 448
 DB 336 K 336

RESULT 3
 AAM50639

ID AAM50639 standard; protein; 336 AA.
AC AAM50639;
XX
DT 04-APR-2002 (first entry)
XX
DE Streptococcus dysgalactiae gapC plasmin binding protein.
XX
KW GapC; plasmin-binding protein; DysgalGapC; infection; mastitis; vaccine;
KW diagnosis; therapy.
XX
OS Streptococcus dysgalactiae.
XX
PN WO2001196381-A2.
XX
PD 20-DEC-2001.
XX
PF 11-JUN-2001; 2001WO-CA000838.
XX
PR 12-JUN-2000; 2000US-0211022P.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
XX
KW WPI; 2002-130725/17.
DR N-PSDB; ABA91248.
XX
XX Novel isolated GapC protein of Streptococcus dysgalactiae, S. agalactiae,
XX S. uberis, S. parvulus, or S. infantis, useful as vaccine component for
XX treating streptococcal infection which causes mastitis in vertebrates.
XX
PS Claim 1(a); Fig 1A-B; 107pp; English.
XX
XX The present sequence is that of the GapC plasmin binding protein
XX (DysgalGapC) of Streptococcus dysgalactiae ATCC 43078, a clinical isolate
XX from bovine mastitis. It is encoded by the GapC gene given in ABA91248.
XX
XX GapC protein, which has no signal sequence or membrane anchor domain, is
XX capable of eliciting an immune response in a vertebrate. The invention
XX provides the GapC genes and proteins of 5 Streptococcus species, as well
XX as recombinant vectors, host cells and vaccine compositions comprising
XX GapC polynucleotides or proteins. The vaccines are used to treat or
XX prevent a bacterial infection, especially a streptococcal infection, and
XX mastitis in particular (claimed). GapC proteins are also used in claimed
XX methods for detecting GapC antibodies, and to raise antibodies that are
XX used in claimed methods for detecting GapC proteins
XX
SQ Sequence 336 AA;

Query Match 72.7%; Score 1656.5; DB 5; Length 336;
Best Local Similarity 79.6%; Pred. No. 4.9e-129;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIQGLAFRIQNVGEVETRIINDLTDPMNLAKLLKYDTTQGRFGDTVEV 87
DB 1 MVKVGINGFGRIQGLAFRIQNVGEVETRIINDLTDPMNLAKLLKYDTTQGRFGDTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKDGFVNGKFTKVSAAE 147
DB 61 KEGGFVNGNFIKVSARDPE----- 81
QY 148 KDPQCIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAREPANIDWATDGVIEVLE 207
DB 82 -----NIDWATDGVIEVLE 95
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDLDGTETVISGASCTTN 267
DB 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDLDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMTLDGHRGGDLRRARAGAAIVPNSTGAA 327
DB 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMTLDGHRGGDLRRARAGAAIVPNSTGAA 215

QY 328 KAIGLIVIPELNGKLDGAAQRPVPTGSVTELWTLDKNVSVDEINAAKKAASNDSPGYTE 387
DB 216 KAIGLIVIPELNGKLDGAAQRPVPTGSVTELWTLDKNVSVDEINAAKKAASNDSPGYTE 275
QY 388 DPVSSDIVGVSYGSLFDATQTKWMEVDGSQLVKVSWYDNMSYTAQLVLTLEYFAKIA 447
DB 276 DPVSSDIVGVSYGSLFDATQTKWMEVDGSQLVKVSWYDNMSYTAQLVLTLEYFAKIA 335
QY 448 K 448
DB 336 K 336

RESULT 4
ABP29960
ID ABP29960 standard; protein; 336 AA.
XX
AC ABP29960;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 9096.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB004789.
XX
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Telford J, Massignani V, Margarit Y RosI, Grandi G, Fraser C;
PI Tetelín H;
XX
DR WPI; 2002-352536/38.
DR N-PSDB; ABN70591.
XX
XX New Streptococcus protein for the treatment or prevention of infection or
XX disease caused by Streptococcus bacteria, such as meningitis, and for
XX detecting a compound that binds to the protein.
XX
PS Claim 1; Page 4031; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
XX streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX the specification. The proteins have antibacterial and antiinflammatory
XX activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
XX antibodies that bind (I) are used in the manufacture of medicaments for
XX the treatment or prevention of infection or disease caused by
XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX Nucleic acids encoding (I) are used to detect Streptococcus in a
XX biological sample. (I) is used to determine whether a compound binds to
XX (I). A composition comprising (I) or a nucleic acid encoding (I), may be
XX used as a vaccine or diagnostic composition. The disease caused by
XX Streptococcus that is prevented or treated may be meningitis. Nucleic
XX acid encoding (I) may be used to recombinantly produce (I) and may be
XX used in gene therapy. Antibodies to (I) are used for affinity
XX chromatography, immunoassays, and distinguishing/identifying
XX Streptococcus proteins
XX
SQ Sequence 336 AA;

for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 74379; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 613 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway; (8) the gene product or that inhibits cellular proliferation; (9) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 336 AA;

Query Match	72.7%;	Score 1655.5;	DB 6;	Length 336;
Best Local Similarity	79.3%;	Pred. No. 5.9e-129;		
Matches 334;	Conservative 1;	Mismatches 1;	Indels 85;	Gaps 1;
QY	28	MVVKVINGFGRIQGLAFRRIONVEGVETRNLDTPNMLAHLKYDTTQGRPDGTVEV	87	
Db	1	MVVKVINGFGRIQGLAFRRIONVEGVETRNLDTPNMLAHLKYDTTQGRPDGTVEV	60	
QY	88	KEGGFEVNGNFIKVSARDPENIDWATDGVIEALEGTVEVKGDFVNGKFIKVS	147	
Db	61	KEGGFEVNGNFIKVSARDPENIDWATDGVIEALEGTVEVKGDFVNGKFIKVS	147	
QY	148	KDPEQIDWATDGVIEALEGTVEVKGDFVNGKFIKVSARDPENIDWATDGVIE	207	
Db	82	-----NIDWATDGVIEALE	95	
QY	208	ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVPNTNHDILDTGTTVSGASCTTN	267	
Db	96	ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVPNTNHDILDTGTTVSGASCTTN	155	
QY	268	CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMLDGPGRGDLRRARAGAAIVPNSTGAA	327	
Db	156	CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMLDGPGRGDLRRARAGAAIVPNSTGAA	215	
QY	328	KAIGLVIPELNGKLDGAAQRPVPTGVTTELVTLDKNVSVDEINAAKASNDSPFGYTE	387	
Db	216	KAIGLVIPELNGKLDGAAQRPVPTGVTTELVTLDKNVSVDEINAAKASNDSPFGYTE	275	
QY	388	DPVSSDIVGVSGLSFDATQTKVMEVDGSQLVKVSWYDNMSYTAQLVRLTLEYFAKIA	447	
Db	276	DPVSSDIVGVSGLSFDATQTKVMEVDGSQLVKVSWYDNMSYTAQLVRLTLEYFAKIA	335	
QY	448	K 448		

448 K 448

Query Match 72.7%; Score 1655.5; DB 5; Length 336;
Best Local Similarity 79.3%; Pred. No. 5.9e-129;
Matches 334; Conservative 1; Mismatches 1; Indels 85; Gaps 1;

QY	28	MVVKVINGFGRIQGLAFRRIONVEGVETRNLDTPNMLAHLKYDTTQGRPDGTVEV	87	
Db	1	MVVKVINGFGRIQGLAFRRIONVEGVETRNLDTPNMLAHLKYDTTQGRPDGTVEV	60	
QY	88	KEGGFEVNGNFIKVSARDPENIDWATDGVIEALEGTVEVKGDFVNGKFIKVS	147	
Db	61	KEGGFEVNGNFIKVSARDPENIDWATDGVIEALEGTVEVKGDFVNGKFIKVS	147	
QY	148	KDPEQIDWATDGVIEALEGTVEVKGDFVNGKFIKVSARDPENIDWATDGVIE	207	
Db	82	-----NIDWATDGVIEALE	95	
QY	208	ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVPNTNHDILDTGTTVSGASCTTN	267	
Db	96	ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVPNTNHDILDTGTTVSGASCTTN	155	
QY	268	CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMLDGPGRGDLRRARAGAAIVPNSTGAA	327	
Db	156	CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMLDGPGRGDLRRARAGAAIVPNSTGAA	215	
QY	328	KAIGLVIPELNGKLDGAAQRPVPTGVTTELVTLDKNVSVDEINAAKASNDSPFGYTE	387	
Db	216	KAIGLVIPELNGKLDGAAQRPVPTGVTTELVTLDKNVSVDEINAAKASNDSPFGYTE	275	
QY	388	DPVSSDIVGVSGLSFDATQTKVMEVDGSQLVKVSWYDNMSYTAQLVRLTLEYFAKIA	447	
Db	276	DPVSSDIVGVSGLSFDATQTKVMEVDGSQLVKVSWYDNMSYTAQLVRLTLEYFAKIA	335	
QY	448	K 448		
Db	336	K 336		

RESULT 5

ABU46455	ABU46455 standard; protein; 336 AA.
XX	ABU46455;
AC	ABU46455;
XX	19-JUN-2003 (first entry)
DT	Protein encoded by Prokaryotic essential gene #31982.
DE	Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX	Streptococcus pyogenes.
XX	WO200277183-A2.
FN	03-OCT-2002.
XX	21-MAR-2002; 2002WO-US009107.
PD	21-MAR-2001; 2001US-00815242.
XX	06-SEP-2001; 2001US-00948993.
PR	25-OCT-2001; 2001US-0342923P.
PR	08-FEB-2002; 2002US-00072851.
PR	06-MAR-2002; 2002US-0362699P.
XX	(ELIT-) ELITRA PHARM INC.
PA	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX	PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
PI	WPI; 2003-029926/02.
XX	DR N-PSDB; ACA50325.
XX	PT New antisense nucleic acids, useful for identifying proteins or screening

Db 336 K 336

RESULT 6
ABP29106
ID ABP29106 standard; protein; 345 AA.

XX AC ABP29106;
XX DT 02-JUL-2002 (first entry)
XX DE Streptococcus polypeptide SEQ ID NO 7388.
XX KW Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;
XX KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX CS Streptococcus pyogenes.
XX PN WO200234771-A2.
XX PD 02-MAY-2002.
XX PF 29-OCT-2001; 2001WO-GB004789.
XX PR 27-OCT-2000; 2000GB-00028333.
XX PR 24-NOV-2000; 2000GB-00028727.
XX PR 07-MAR-2001; 2001GB-00005640.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX PI Telford J, Massignani V, Margarit Y Rosi, Grandi G, Fraser C;
XX PI Tettelin H;
XX XX WPI; 2002-352536/38.
XX DR N-PSDB; ABN69737.
XX PT New Streptococcus protein for the treatment or prevention of infection or
XX PT disease caused by Streptococcus bacteria, such as meningitis, and for
XX PT detecting a compound that binds to the protein.
XX PS Claim 1; Page 388; 4525pp; English.
XX CC The invention relates to a protein (ABP25413-ABP30895) from group B
XX CC Streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX CC the specification. The proteins have antibacterial and antiinflammatory
XX CC activity. (I), nucleic acids encoding (I), ABN65044-ABN71526 and
XX CC antibodies that bind (I) are used in the manufacture of medicaments for
XX CC the treatment or prevention of infection or disease caused by
XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX CC Nucleic acids encoding (I) are used to detect Streptococcus in a
XX CC biological sample. (I) is used to determine whether a compound binds to
XX CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
XX CC used as a vaccine or diagnostic composition. The disease caused by
XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic
XX CC acid encoding (I) may be used to recombinantly produce (I) and may be
XX CC used in gene therapy. Antibodies to (I) are used for affinity
XX CC chromatography, immunoassays, and distinguishing/identifying
XX CC Streptococcus proteins
XX SQ Sequence 345 AA;

Query Match 72.7%; Score 1655.5; DB 5; Length 345;
Best Local Similarity 79.3%; Pred. No. 6.1e-129;
Matches 334; Conservative 1; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGFRIGLAFRRIONVEGVETRLNDLDPNLAHLKYDTTQGRFDGTVEV 87
DB 10 MVVKVINGFGFRIGLAFRRIONVEGVETRLNDLDPNLAHLKYDTTQGRFDGTVEV 69
QY 88 KEGGFVNGNFIKVSAREDPENIDWATDGEIVLEALEGTVVKDGGDFVNGKFIKVAE 147

Db 70 KEGGFVNGNFIKVSAREDPENIDWATDGEIVLEALEGTVVKDGGDFVNGKFIKVAE 90
QY 148 KDPQIDWATDGEIVLEIDGTVEYKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207
Db 91 -----NIDWATDGEIVLE 104
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTTETVSGASCTTN 267
Db 105 ATGFFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVNTNHDILDGTTETVSGASCTTN 164
QY 268 CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMLDGPGRGDLRRARAGAANIVPNSTGAA 327
Db 165 CLAPMAKALHDAFGIQKGLMTTTHAYTGDQMLDGPGRGDLRRARAGAANIVPNSTGAA 224
QY 328 KAIGLVIPELANGKLDGAAQRPVPTGVTETLVTLDRKVSVDENAKMAASNDVSFGYTE 387
Db 225 KAIGLVIPELANGKLDGAAQRPVPTGVTETLVTLDRKVSVDENAKMAASNDVSFGYTE 284
QY 388 DPVSSDIVGVSYGSLFDATQTKWNEVDGSLVKVSWYDNEMSVTAQLVTTLEYFAKIA 447
Db 285 DPVSSDIVGVSYGSLFDATQTKWNEVDGSLVKVSWYDNEMSVTAQLVTTLEYFAKIA 344
QY 448 K 448
Db 345 K 345

RESULT 7

AAR56486
ID AAR56486 standard; protein; 336 AA.

XX AC AAR56486;
XX DT 25-MAR-2003 (revised)
XX DT 22-FEB-1995 (first entry)
XX DE Plasmin receptor.
XX KW Plasmin receptor; isolate 64/14; plasmin; alpha-2-antiplasmin;
XX KW plasminogen activator; bleeding; reocclusion; thrombosis;
XX KW pulmonary embolism; clots.
XX OS Streptococcus pyogenes.

XX US3328996-A.

XX 12-JUL-1994.

XX 10-AUG-1992; 92US-00928462.

XX 29-MAR-1989; 89US-00330849.

XX 16-MAY-1990; 90US-00524411.

XX (UYFL) UNIV FLORIDA RES FOUND INC.

XX Von Mering G, Broder C, Boyle MDP, Lottenberg R;

XX WPI; 1994-225327/27.

XX N-PSDB; AAQ70705.

XX New DNA encoding bacterial plasmin receptor - useful as thrombolytic
XX agents, used with plasminogen activator or bound to plasmin, also useful
XX in vaccines against bacterial infection.

XX Claim 1; Col 27-30; 19pp; English.

XX This sequence represents the S. pyogenes plasmin receptor. The DNA
XX encoding this sequence was isolated from the S. pyogenes clinical isolate
XX 64/14. The plasmin receptor has a very high affinity for plasmin which,
XX when bound, retains its enzymatic activity but is not regulated
XX (inhibited) by alpha-2-antiplasmin. The receptor protein, when
XX administered concurrently or sequentially, prolongs the activity of

CC plasminogen activator (PA) so allows a reduction in dose, and thus lowers
 CC the risk of bleeding, and may prevent reocclusion of blood vessels. The
 CC protein may be coupled to a fibrin-specific monoclonal antibody to
 CC provide targeting to clots. The plasmin receptor may be useful in human
 CC or veterinary medicine, for treatment of thrombosis and pulmonary
 CC embolism, and for solubilizing clots in catheters or shunts. (Updated on
 CC 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 336 AA;

Query Match 72.5%; Score 1652.5; DB 2; Length 336;
 Best Local Similarity 79.1%; Pred. No. 1e-128;
 Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;
 QY 28 MVKVGINGFGRIQLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRPDGTVEV 87
 DB 1 MVKVGINGFGRIQLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRPDGTVEV 60
 QY 88 KEGGFVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVDGGFDVNGKFIKVS 147
 DB 61 KEGGFVNGNFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFVNGQFVKVSAREPANIDWATDGEIVLE 207
 DB 82 -----NIDWATDGEIVLE 95
 QY 208 ATSPFAKKEAAEKHLHANGAKKVITAPGNDVKTVVFNTHDILDGTETVIGASCTTN 267
 DB 96 ATGFFAKKEAAEKHLHANGAKKVITAPGNDVKTVVFNTHDILDGTETVIGASCTTN 155
 QY 268 CLAPMAKALHDAPFGIQKGLMTTHAYTGDMILDPHRRGGLRRARAGAAIVPNSTGAA 327
 DB 156 CLAPMAKALHDAPFGIQKGLMTTHAYTGDMILDPHRRGGLRRARAGAAIVPNSTGAA 215
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVLDKNVSVDEINAAKAAASNDSPGYTE 387
 DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVLDKNVSVDEINAAKAAASNDSPGYTE 275
 QY 388 DPVSSDIVGVSIGSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
 DB 276 DPVSSDIVGVSIGSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
 QY 448 K 448
 DB 336 K 336

RESULT 8
 AAY85681
 ID AAY85681 standard; protein; 336 AA.
 AC AAY85681;
 XX
 DT 13-FEB-2001 (first entry)
 XX
 DE Streptococcal plasmin receptor amino acid sequence.
 XX
 KW Plasmin receptor; plr; immune response; vaccination; antibacterial;
 KW Streptococcal infection; antihelminthic.
 XX
 OS Streptococcus pyogenes.
 XX
 PN US6136323-A.
 XX
 PD 24-OCT-2000.
 XX
 PF 11-JUL-1994; 94US-00273247.
 XX
 PR 29-MAR-1989; 89US-00330849.
 PR 16-MAY-1990; 90US-00524411.
 PR 10-AUG-1992; 92US-00928462.
 XX
 PA (UYFL) UNIV FLORIDA RES FOUND INC.

XX Von Mering G, Broder C, Lottenberg R, Boyle MDP;
 XX WPI; 2001-006210/01.
 DR N-P8DB; AAC66140.
 XX
 PT Raising an immune response in a mammal, especially for vaccination
 PT against group A streptococcal infections, comprises administering an
 PT isolated streptococcal plasmin receptor protein.
 XX
 PS Claim 2; Col 27-30; 17pp; English.
 XX
 CC Polynucleotide sequence AAC66140 encodes the Streptococcal plasmin
 CC receptor protein AAY85681. The protein is used in a method for raising an
 CC immune response in a mammal. The method comprises administering the
 CC plasmin receptor protein. The method is useful as a vaccination against
 CC group A Streptococcal infections and potentially against a broad range of
 CC infections associated with pathogens expressing glyceraldehyde-3-
 CC phosphate dehydrogenase (GAPDH) or GAPDH-like proteins at their cell
 CC surfaces. The vaccine has antibacterial and antihelminthic activity
 XX
 SQ Sequence 336 AA;

Query Match 72.5%; Score 1652.5; DB 4; Length 336;
 Best Local Similarity 79.1%; Pred. No. 1e-128;
 Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;
 QY 28 MVKVGINGFGRIQLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRPDGTVEV 87
 DB 1 MVKVGINGFGRIQLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRPDGTVEV 60
 QY 88 KEGGFVNGNFIKVSARDPENIDWATDGEIVLEALEGTVEVDGGFDVNGKFIKVS 147
 DB 61 KEGGFVNGNFIKVSARDPE----- 81
 QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGGFVNGQFVKVSAREPANIDWATDGEIVLE 207
 DB 82 -----NIDWATDGEIVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVITAPGNDVKTVVFNTHDILDGTETVIGASCTTN 267
 DB 96 ATGFFAKKEAAEKHLHANGAKKVITAPGNDVKTVVFNTHDILDGTETVIGASCTTN 155
 QY 268 CLAPMAKALHDAPFGIQKGLMTTHAYTGDMILDPHRRGGLRRARAGAAIVPNSTGAA 327
 DB 156 CLAPMAKALHDAPFGIQKGLMTTHAYTGDMILDPHRRGGLRRARAGAAIVPNSTGAA 215
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVLDKNVSVDEINAAKAAASNDSPGYTE 387
 DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVTVELVLDKNVSVDEINAAKAAASNDSPGYTE 275
 QY 388 DPVSSDIVGVSIGSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
 DB 276 DPVSSDIVGVSIGSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
 QY 448 K 448
 DB 336 K 336

RESULT 9
 ABP30758
 ID ABP30758 standard; protein; 336 AA.
 XX
 AC ABP30758;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 10692.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX OS Streptococcus agalactiae.
XX PN WO200234771-A2.
XX XX
XX PD 02-MAY-2002.
XX PF 29-OCT-2001; 2001WO-GB004789.
XX PR 27-OCT-2000; 2000GB-00026333.
XX PR 24-NOV-2000; 2000GB-00028727.
XX PR 07-MAR-2001; 2001GB-00005640.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX PI Telford J, Masignani V, Margarit Y Rosl, Grandi G, Fraser C;
XX PI Tettelin H;
XX DR N-PSDB; ABN71389.
XX DR WPI; 2002-352536/38.
XX DR N-PSDB; ABN71389.
XX XX
XX PT New Streptococcus protein for the treatment or prevention of infection or
XX PT disease caused by Streptococcus bacteria, such as meningitis, and for
XX PT detecting a compound that binds to the protein.
XX PS Claim 1; Page 4179; 4525pp; English.
XX XX
XX CC The invention relates to a protein (ABP25413-ABP30895) from group B
XX CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX CC the specification. The proteins have antibacterial and antiinflammatory
XX CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
XX CC antibodies that bind (I) are used in the manufacture of medicaments for
XX CC the treatment or prevention of infection or disease caused by
XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX CC Nucleic acids encoding (I) are used to detect Streptococcus in a
XX CC biological sample. (I) is used to determine whether a compound binds to
XX CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
XX CC used as a vaccine or diagnostic composition. The disease caused by
XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic
XX CC acid encoding (I) may be used to recombinantly produce (I) and may be
XX CC used in gene therapy. Antibodies to (I) are used for affinity
XX CC chromatography, immunoassays, and distinguishing/identifying
XX CC Streptococcus proteins
XX XX
SQ Sequence 336 AA;
Query Match 68.7%; Score 1564.5; DB 5; Length 336;
Best Local Similarity 74.1%; Pred. No. 2.1e-121;
Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;
QY 28 MVKVGINGFGRIGRLAFRIQNVGVETVRINDLTPNMLAHLKYDTTQGRFPGTVEV 87
DB 1 MVKVGINGFGRIGRLAFRIQNVGVETVRINDLTPNMLAHLKYDTTQGRF----- 54
QY 88 KEGGEVWGNFKVSAERDPENIDWATDGVETVLEALETGVKDGVDVNGKFKVSAE 147
DB 55 ----- 54
QY 148 KDPEQIDWATDGVETVLEIDGTVEKGGFEVNGQFVKVSAEREPANIDWATDGVETVLE 207
DB 55 -----DGTEVEKGGFEVNGQFVKVSAEREPANIDWATDGVETVLE 95
QY 208 ATSPFAKKEAARKHLHANGAKKVVITAPGNDVKVTVFNTNHDLDGTETVISGASCTTN 267
DB 96 ATGFFASKEBAKQKHENGAKKVVITAPGNDVKVTVFNTNHDLDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDFAFGIQKGLMTTHAYTQDMLDGFHGGDLRRRAGAANIVFNSTGAA 327
DB 156 CLAPMAKALQDNFGVKQGLMTTHAYTQDMLDGFHGGDLRRRAGAANIVFNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAARQVPVPTGVTVELVTLTKDNVSDVEINAAKAAANSDFGYTE 387

DB 216 KAIGLVIPELNGKLDGAARQVPVPTGVTVELVTLTKDNVSDVEINAAKAAANSDFGYTE 275
QY 388 DPIVSSDIVGVSGSLFDATQTKMVEVDGSQLVKVSWYDNMSYTAQLVRLTLEYFAKIA 447
DB 276 DPIVSSDIVGVSGSLFDATQTKVQTVGDNQNLVKVSWYDNMSYTSQVRLTLEYFAKIA 335
QY 448 K 448
DB 336 K 336
RESULT 10
AAM50667
ID AAM50667 standard; protein; 336 AA.
XX XX
AC AAM50667;
XX XX
DT 08-APR-2002 (first entry)
XX XX
DE Streptococcus uberis gapC plasmin binding protein UberGapC.
XX XX
KW UberGapC; GapC; plasmin binding protein; epitope; infection; vaccine;
XX KW immunisation; mastitis; therapy.
XX OS Streptococcus uberis.
XX PN WO200196379-A2.
XX XX
PD 20-DEC-2001.
XX PF 11-JUN-2001; 2001WO-CA000836.
XX PR 12-JUN-2000; 2000US-0211247P.
XX PA (UYSA-) UNIV SASKATCHEWAN.
XX PI Potter AA, Perez-Casal J, Fontaine M;
XX PI WPI; 2002-098051/13.
XX DR N-PSDB; ABA91330.
XX XX
PT Novel GapC multiple epitope fusion polypeptide comprising antigenic
PT determinant of Streptococcus dysgalactiae S. agalactiae, S. uberis, S.
PT parauberis, or S. iniae GapC protein useful for treating mastitis in
PT vertebrates.
XX XX
PS Claim 8; Fig 3A-B; 116pp; English.
XX XX
CC The present sequence is that of the GapC plasmin binding protein.
CC UberGapC of Streptococcus uberis ATCC 927. The invention relates to
CC novel GapC multiple epitope fusion proteins that comprise epitopes from 1
CC or more of Streptococcus dysgalactiae, Streptococcus agalactiae,
CC Streptococcus parauberis, Streptococcus uberis and Streptococcus iniae
CC (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664).
CC Expression vectors and host cells for production of the multiple epitope
CC fusion proteins are provided. The multiple epitope proteins are used in
CC claimed vaccines for treating or preventing a bacterial infection in a
CC vertebrate, especially a streptococcal infection, and particularly
CC mastitis. They are also used in claimed methods of detecting
CC Streptococcus antibodies. The multiple epitope proteins are capable of
CC eliciting broad immunity against a variety of streptococcal infections
CC while minimising the number of antigens present in the final formulation
CC and concomitantly reducing production costs
XX XX
SQ Sequence 336 AA;
Query Match 68.5%; Score 1559.5; DB 5; Length 336;
Best Local Similarity 73.9%; Pred. No. 5.5e-121;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVKVGINGFGRIGRLAFRIQNVGVETVRINDLTPNMLAHLKYDTTQGRFPGTVEV 87

Db 1 MVKVGINGFRIGRLAFRIQNVGVETRIINDLTPNMLHLKYDTTQGRFDGTEVEV 60
 QY 88 KEGGFEVNGNFIKVAERDPENIDWATDGEIVLEALEGTVEVKDGGDFVNGKFIKVSAAE 147
 Db 61 KGGFEVNGNFIKVAERDPE----- 81
 QY 148 KDPEQIDWATDGEIVLEIDGIVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207
 Db 82 -----NIDWATDGEIVLE 95
 QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTGTEVVISGASCTTN 267
 Db 96 ATGFFAKKAAAEKHLHANGAKKVVITAPGGDDVKTVVFNTHDILDTGTEVVISGASCTTN 155
 QY 268 CLAPMAKALHDFAFGIOKGLMTTHAYTGDQMLDGPGRGDLRRARAGASNIVPNSTGAA 327
 Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDQMLDGPGRGDLRRARAGASNIVPNSTGAA 215
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTVTLVTLTKNVSVDENIINAAKAAASNDSEFGYTE 387
 Db 216 KAIGLVIPELNGKLDGAAQRPVPTGTVTLVTLVAVLEKETSVEEINAAKAAASNDSEFGYTE 275
 QY 388 DPVSSDIIVGSGSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
 Db 276 DPVSSDIIGMAYGSLFDATQTKVQTVVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
 QY 448 K 448
 Db 336 K 336

RESULT 11

AAM50641
 ID AAM50641 standard; protein; 336 AA.

XX AC AAM50641;
 XX DT 04-APR-2002 (first entry)
 XX DE Streptococcus uberis gapC plasmin binding protein.
 XX KW GapC; plasmin-binding protein; UberGapC; infection; mastitis; vaccine;
 XX OS diagnosis; therapy.
 XX OS Streptococcus uberis.

XX PN WO200196381-A2.

XX PD 20-DEC-2001.

XX PF 11-JUN-2001; 2001WO-CA000838.

XX PR 12-JUN-2000; 2000US-0211022P.

XX PA (UYSA-) UNIV SASKATCHEWAN.

XX PI Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;

XX DR WPI; 2002-130725/17.

XX DR N-PSDB; ABA91250.

XX PT Novel isolated GapC protein of Streptococcus dysgalactiae, S.agalactiae,
 PT S.uberis, S.paraberis, or S.iniae, useful as vaccine component for
 PT treating streptococcal infection which causes mastitis in vertebrates.

XX PS Claim 1(c); Fig 3A-B; 107pp; English.

XX CC The present sequence is that of the GapC plasmin binding protein
 CC (UberGapC) of Streptococcus uberis ATCC 9927, encoded by the GapC gene
 CC given in ABA91250. GapC, which has no signal sequence or membrane anchor
 CC domain, is capable of eliciting an immune response in a vertebrate. The
 CC invention provides the GapC genes and proteins of 5 Streptococcus
 CC species, as well as recombinant vectors, host cells and vaccine

CC compositions comprising GapC polynucleotides or proteins. The vaccines
 CC are used to treat or prevent a bacterial infection, especially a
 CC streptococcal infection, and mastitis in particular (claimed). GapC
 CC proteins are also used in claimed methods for detecting GapC antibodies,
 CC and to raise antibodies that are used in claimed methods for detecting
 CC GapC proteins. S. uberis is a common pathogen associated with mastitis in
 CC cattle, horse, sheep and goat
 XX

XX Sequence 336 AA;

Query Match 68.5%; Score 1559.5; DB 5; Length 336;
 Best Local Similarity 73.9%; Pred. No. 5.5e-121;
 Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVKVGINGFRIGRLAFRIQNVGVETRIINDLTPNMLHLKYDTTQGRFDGTEVEV 87

Db 1 MVKVGINGFRIGRLAFRIQNVGVETRIINDLTPNMLHLKYDTTQGRFDGTEVEV 60

QY 88 KEGGFEVNGNFIKVAERDPENIDWATDGEIVLEALEGTVEVKDGGDFVNGKFIKVSAAE 147

Db 61 KGGFEVNGNFIKVAERDPE----- 81

QY 148 KDPEQIDWATDGEIVLEIDGIVEVKEGGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207

Db 82 -----NIDWATDGEIVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDTGTEVVISGASCTTN 267

Db 96 ATGFFAKKAAAEKHLHANGAKKVVITAPGGDDVKTVVFNTHDILDTGTEVVISGASCTTN 155

QY 268 CLAPMAKALHDFAFGIOKGLMTTHAYTGDQMLDGPGRGDLRRARAGASNIVPNSTGAA 327

Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDQMLDGPGRGDLRRARAGASNIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTVTLVTLTKNVSVDENIINAAKAAASNDSEFGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQRPVPTGTVTLVTLVAVLEKETSVEEINAAKAAASNDSEFGYTE 275

QY 388 DPVSSDIIVGSGSLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447

Db 276 DPVSSDIIGMAYGSLFDATQTKVQTVVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 12

AAM50666

ID AAM50666 standard; protein; 336 AA.

XX AC AAM50666;

XX DT 08-APR-2002 (first entry)

XX DE Streptococcus agalactiae gapC plasmin binding protein AgalGapC.

XX KW AgalGapC; GapC; plasmin binding protein; epitope; infection; vaccine;
 XX immunisation; mastitis; therapy.

XX OS Streptococcus agalactiae.

XX PN WO200196379-A2.

XX PD 20-DEC-2001.

XX PF 11-JUN-2001; 2001WO-CA000836.

XX PR 12-JUN-2000; 2000US-0211247P.

XX PA (UYSA-) UNIV SASKATCHEWAN.

XX PI Potter AA, Perez-Casal J, Fontaine M;

XX WPI; 2002-098051/13.
DR N-PSDB; ABA91329.
XX Novel GapC multiple epitope fusion polypeptide comprising antigenic
PT determinant of Streptococcus dysgalactiae, S. agalactiae, S. uberis, S.
PT parauberis, or S. iniae GapC protein useful for treating mastitis in
PT vertebrates.
XX Claim 8; Fig 2A-B; 116pp; English.
PS
XX
XX The present sequence is that of the GapC plasmin binding protein,
CC AgalGapC, of Streptococcus agalactiae ATCC 27541. The invention relates
CC to novel GapC multiple epitope fusion proteins that comprise epitopes
CC from 1 or more of Streptococcus dysgalactiae, Streptococcus agalactiae,
CC Streptococcus parauberis, Streptococcus uberis and Streptococcus iniae
CC (see AAM50665-69). A claimed example is PolyGap4 (see AAM50664).
CC Expression vectors and host cells for production of the multiple epitope
CC fusion proteins are provided. The multiple epitope proteins are used in
CC claimed vaccines for treating or preventing a bacterial infection in a
CC vertebrate, especially a streptococcal infection, and particularly
CC mastitis. They are also used in claimed methods of detecting
CC Streptococcus antibodies. The multiple epitope proteins are capable of
CC eliciting broad immunity against a variety of streptococcal infections
CC while minimising the number of antigens present in the final formulation
CC and concomitantly reducing production costs
XX
SQ Sequence 336 AA;
Query Match 68.4%; Score 1557.5; DB 5; Length 336;
Best Local Similarity 73.9%; Pred. No. 8.1e-121;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVYKVGINGFRIGRLAFRRIQNVEGVETVRINDLDPNMLAHLKYDTTQGRFDGTVEV 87
DB 1 MVYKVGINGFRIGRLAFRRIQNVEGVETVRINDLDPNMLAHLKYDTTQGRF----- 54
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEAEGTVKVGDFVNGKFIKVS 147
DB 55 ----- 54
QY 148 KDEQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAREPANIDWATDGVIEVLE 207
DB 55 -----DGTVEVKEGGFVNGQFVKVSAREPANIDWATDGVIEVLE 95
QY 208 ATSPFAKKEAEKHUANGAKKVVITAPGGNDVKVTWNTNHDILDTGTETVISGASCTTN 267
DB 96 ATGFFASKEKAGQHIHENGAKKVVITAPGGNDVKVTWNTNHDILDTGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAGFQKGLMTTTHAYTGDMILDPHGRGDLRRARAGAAIVPNSTGAA 327
DB 156 CLAPMAKALQDNFVGKQGLMTTTHAYTGDMILDPHGRGDLRRARAGAAIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLKNSVDENNAKAAANDSFYGYTE 387
DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLKNSVDENNAKAAANDSFYGYTE 275
QY 388 DPVSSDIVGVSYGLFATQTKNVEVDSGLVKVSVYDNEMSVYTAOLVTLVTFYAKIA 447
DB 276 DPVSSDIVGVSYGLFATQTKNVEVDSGLVKVSVYDNEMSVYTAOLVTLVTFYAKIA 335
QY 448 K 448
DB 336 K 336
RESULT 13
ID AAM50640
XX AAM50640 standard; protein; 336 AA.
XX
AC AAM50640;
XX
DT 04-APR-2002 (first entry)

XX Streptococcus agalactiae gapC plasmin binding protein.
DE
XX GapC; plasmin-binding protein; AgalGapC; infection; mastitis; vaccine;
KW diagnosis; therapy.
XX Streptococcus agalactiae.
OS
PN WO2001196381-A2.
XX
PD 20-DEC-2001.
XX
XX 11-JUN-2001; 2001WO-CA000838.
PF
XX 12-JUN-2000; 2000US-0211022P.
XX
XX (UYSA-) UNIV SASKATCHEWAN.
PA
XX Bolton AJ, Perez-Casal J, Fontaine M, Potter AA;
PI WPI; 2002-130725/17.
DR N-PSDB; ABA91249.
XX Novel isolated GapC protein of Streptococcus dysgalactiae, S. agalactiae,
PT S. uberis, S. parauberis, or S. iniae, useful as vaccine component for
PT treating streptococcal infection which causes mastitis in vertebrates.
PT
XX Claim 1(b); Fig 2A-B; 107pp; English.
PS
XX The present sequence is that of the GapC plasmin binding protein
CC (AgalGapC) of Streptococcus agalactiae ATCC 27541, encoded by the GapC
CC gene given in ABA91249. GapC, which has no signal sequence or membrane
CC anchor domain, is capable of eliciting an immune response in a
CC vertebrate. The invention provides the GapC genes and proteins of 5
CC Streptococcus species, as well as recombinant vectors, host cells and
CC vaccine compositions comprising GapC polynucleotides or proteins. The
CC vaccines are used to treat or prevent a bacterial infection, especially a
CC streptococcal infection, and mastitis in particular (claimed). GapC
CC proteins are also used in claimed methods for detecting GapC antibodies,
CC and to raise antibodies that are used in claimed methods for detecting
CC GapC proteins. S. agalactiae is a common pathogen associated with
CC mastitis in cattle, horse, sheep and goat. It also causes septicemia,
CC meningitis, bacteraemia, impetigo, arthritis, urinary tract infections,
CC abscesses, spontaneous abortion, etc
XX
SQ Sequence 336 AA;
Query Match 68.4%; Score 1557.5; DB 5; Length 336;
Best Local Similarity 73.9%; Pred. No. 8.1e-121;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVYKVGINGFRIGRLAFRRIQNVEGVETVRINDLDPNMLAHLKYDTTQGRFDGTVEV 87
DB 1 MVYKVGINGFRIGRLAFRRIQNVEGVETVRINDLDPNMLAHLKYDTTQGRF----- 54
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEAEGTVKVGDFVNGKFIKVS 147
DB 55 ----- 54
QY 148 KDEQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAREPANIDWATDGVIEVLE 207
DB 55 -----DGTVEVKEGGFVNGQFVKVSAREPANIDWATDGVIEVLE 95
QY 208 ATSPFAKKEAEKHUANGAKKVVITAPGGNDVKVTWNTNHDILDTGTETVISGASCTTN 267
DB 96 ATGFFASKEKAGQHIHENGAKKVVITAPGGNDVKVTWNTNHDILDTGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAGFQKGLMTTTHAYTGDMILDPHGRGDLRRARAGAAIVPNSTGAA 327
DB 156 CLAPMAKALQDNFVGKQGLMTTTHAYTGDMILDPHGRGDLRRARAGAAIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLKNSVDENNAKAAANDSFYGYTE 387
DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLKNSVDENNAKAAANDSFYGYTE 275

CC media or ear infection. They are also useful in developing vaccines,
 CC diagnostics and antibiotics. The methods are useful for identifying
 CC immunodominant proteins. The present sequence is one of the 2469 proteins
 CC expressed by the identified coding regions from the genomic sequence.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 335 AA;
 Query Match 67.4%; Score 1535; DB 6; Length 335;
 Best Local Similarity 73.6%; Pred. No. 6e-119;
 Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;
 QY 28 MVKVGINGFGRIGRLAFRIQNVGVETRIINDLTPNMLAHLLKXDTTQGRFDGTEV 87
 DB 1 MVKVGINGFGRIGRLAFRIQNVGVETRIINDLTPNMLAHLLKXDTTQGRFDGTEV 60
 QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGFKIVSAE 147
 DB 61 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGFKIVSAE 81
 QY 148 KDPQIDWATDGVLEIVLEIDGTVEVKEGPFVNGQFVKVSAEREPANIDWATDGVLEIVLE 207
 DB 82 -----QIDWATDGVLEIVLE 95
 QY 208 ATSPFAKKEAEAKULHANGAKKVVITAPGNDVKTVVNTNHDILDTETVISCASCTTN 267
 DB 96 ATGFFAKKEAEAKULHANGAKKVVITAPGNDVKTVVNTNHDILDTETVISCASCTTN 154
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANVPSNCGAA 327
 DB 155 CLAPMAKALQDNFVGVEGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANVPSNCGAA 214
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVWLDKNSVDEINAAKASNDSPGVTE 387
 DB 215 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVWLDKNSVDEINAAKASNDSPGVTE 274
 QY 388 DPVSSDIVGVSGLFDATQTKVNEVDGSQLVKVSWYDNEMSYTAQLVTLTYFAKIA 447
 DB 275 DPVSSDIVGVSGLFDATQTKVNEVDGSQLVKVSWYDNEMSYTAQLVTLTYFAKIA 334
 QY 448 K 448
 DB 335 K 335
 RESULT 15
 AAU37576
 ID AAU37576 standard; protein; 359 AA.
 XX
 AC AAU37576;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Streptococcus pneumoniae cellular proliferation protein #5.
 XX
 KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
 XX antibacterial; drug design.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2001; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.

DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVWLDKNSVDEINAAKASNDSPGVTE 275
 QY 388 DPVSSDIVGVSGLFDATQTKVNEVDGSQLVKVSWYDNEMSYTAQLVTLTYFAKIA 447
 DB 275 DPVSSDIVGVSGLFDATQTKVNEVDGSQLVKVSWYDNEMSYTAQLVTLTYFAKIA 335
 QY 448 K 448
 DB 335 K 335
 RESULT 14
 ABU02516
 ID ABU02516 standard; protein; 335 AA.
 XX
 AC ABU02516;
 XX
 DT 23-OCT-2003 (revised)
 DT 11-FEB-2003 (first entry)
 XX
 DE S. pneumoniae type 4 strain protein from coding region #2094.
 XX
 KW Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
 KW anti-inflammatory; antibacterial; immunostimulant; auditory; respiratory;
 KW gene therapy; vaccine.
 XX
 OS Streptococcus pneumoniae; type 4 strain.
 XX
 PN WO200277021-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 27-MAR-2002; 2002WO-IB002163.
 XX
 PR 27-MAR-2001; 2001GB-00007658.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Masignani V, Tettelin H, Fraser C;
 XX
 DR WPI; 2003-040579/03.
 DR N-PSDB; ABX07806.
 XX
 PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
 PT ear infection.
 XX
 PS Claim 1; SEQ ID NO 4188; 56pp; English.
 XX
 CC The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC ABS56454. Also included are an antibody which binds one of the proteins,
 CC treating a patient by administering the protein, DNA or antibody (in a
 CC composition), a kit comprising first and second primers, which are the
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a
 CC sequence not defined in the specification, for amplifying a target
 CC sequence contained within a Streptococcus nucleic acid sequence, where
 CC the first primer is substantially complementary to the target sequence
 CC and the second primer is substantially complementary to the complement of
 CC the target sequence, and where the parts of the primers having
 CC substantial complementarity define the termini of the target sequence to
 CC be amplified, assay comprising contacting a test compound with the
 CC protein, and determining whether the test compound binds to the protein
 CC and a Streptococcus pneumoniae bacterium, where one or more genes
 CC encoding the proteins has been rendered inactive. The proteins, nucleic
 CC acid molecules, antibody and compositions are useful as medicaments for
 CC treating or preventing a disease or infection due to streptococcus
 CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis

Search completed: March 29, 2004, 16:33:20
Job time : 99.4776 secs

```
PR 27-NOV-2000; 2000US-0253625P.  
PR 22-DEC-2000; 2000US-0257931P.  
PR 16-FEB-2001; 2001US-0269308P.  
XX  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HH;  
XX  
XX WPI; 2001-611495/70.  
DR N-PSDB; AAS55435.  
XX  
XX New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids.  
PT  
XX  
PS Example 3; SEQ ID NO 13169; 511pp; English.  
XX  
XX The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the genes,  
CC their use in the discovery of novel antibiotics, the essential genes  
CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
CC useful for the identification of potential new targets for antibiotic  
CC development. The antisense nucleic acids can also be used to identify  
CC proteins used in proliferation, to express these proteins, and to obtain  
CC antibodies capable of binding to the expressed proteins. The proteins can  
CC be used to screen compounds in rational drug discovery programmes. The  
CC antisense nucleic acid sequence is also useful to screen for homologous  
CC nucleic acids which are required for cell proliferation in a wide variety  
CC of organisms. The present sequence represents an essential prokaryotic  
CC cellular proliferation protein. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published_pct_sequences  
XX  
XX SQ Sequence 359 AA;  
  
Query Match 67.4%; Score 1535; DB 4; Length 359;  
Best Local Similarity 73.6%; Pred. No. 6.6e-119;  
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;  
  
QY 28 MVVKYINGFGRIGLAFRRIONVEGVETRIINDLTDPNLAHLKDYDTQGRFDGTVEV 87  
Db |||||  
QY 25 MVVKYINGFGRIGLAFRRIONVEGVETRIINDLTDPNLAHLKDYDTQGRFDGTVEV 84  
Db |||||  
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147  
Db |||||  
QY 85 KEGGFVNGKFIKVSARDPE----- 105  
QY 148 KDPEQIDWATDGVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVLEIVLE 207  
Db |||||  
QY 106 -----QIDWATDGVLEIVLE 119  
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTWVNTNHDILDGTETVISGASCTTN 267  
Db |||||  
QY 120 ATGFFAKKEAAEKHL-KGAKKVVITAPGNDVKTWVNTNHDILDGTETVISGASCTTN 178  
QY 268 CLAPMAKALHDAFGIQKGLMTTHIAYTGDQIMLDGPHRGGDLRRARAGAAIVPNSTGAA 327  
Db |||||  
QY 179 CLAPMAKALQDNFGVYVEGLMTTHIAYTGDQIMLDGPHRGGDLRRARAGAAIVPNSTGAA 238  
QY 328 KAIGLVIPELNGKLDGAQRVPVPTGVTVELVWTLDKVSVDEINAAKASNDSPGYTE 387  
Db |||||  
QY 239 KAIGLVIPELNGKLDGSAQRVPVPTGVTVELVWTLDEVNAAKASNESYGYTE 298  
QY 388 DFISSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVETLEYFAKIA 447  
Db |||||  
QY 299 DFISSDIVGVSGLFDATQTKVLDVDGKQLVKVSWYDNEMSYTAQLVETLEYFAKIA 358  
QY 448 K 448  
Db 359 K 359
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 16:31:39 ; Search time 31.5224 Seconds
(without alignments)
733.715 Million cell updates/sec

Title: US-10-650-369-22

Perfect score: 2278

Sequence: 1 MKKITGILLALLAVILSAC.....EMSYAQLVRLTFYFAKIAK 448

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2278	100.0	448	4	US-09-878-766A-22
2	1656.5	72.7	336	4	US-09-878-766A-12
3	1652.5	72.5	336	1	US-07-928-462-2
4	1652.5	72.5	336	3	US-08-273-247-2
5	1559.5	68.5	336	4	US-09-878-766A-16
6	1557.5	68.4	336	4	US-09-878-766A-14
7	1534.5	67.4	336	4	US-09-878-766A-20
8	1530.5	67.2	336	4	US-09-878-766A-18
9	1491	65.5	333	3	US-08-961-083-54
10	1491	65.5	333	4	US-09-536-784-54
11	1282	56.3	346	4	US-09-134-000C-4400
12	1270	55.8	333	4	US-09-107-532A-4769
13	1138.5	50.0	348	4	US-09-134-001C-5513
14	923.5	40.5	357	4	US-09-134-000C-4229
15	918.5	40.3	340	4	US-09-634-238-233
16	905.5	39.7	349	4	US-09-489-039A-13939
17	849	37.3	335	6	5290690-9
18	838	36.8	334	6	5290690-10
19	803	35.3	340	3	US-09-095-855-205
20	803	35.3	340	4	US-09-205-428-205
21	780	34.2	334	6	5290690-11
22	776.5	34.1	335	4	US-09-489-039A-7679
23	758.5	33.3	328	4	US-09-107-532A-4057
24	747.5	32.8	340	4	US-09-543-681A-8083
25	719	31.6	352	4	US-09-489-039A-9410
26	690	30.3	336	4	US-09-198-452A-662
27	683.5	30.0	338	4	US-09-091-725-10

28	678	29.8	268	2	US-08-997-080-186	Sequence 186, App
29	678	29.8	268	3	US-08-997-362-186	Sequence 186, App
30	678	29.8	268	3	US-09-095-855-186	Sequence 186, App
31	678	29.8	268	4	US-09-324-542-186	Sequence 186, App
32	678	29.8	268	4	US-09-205-428-186	Sequence 186, App
33	675	29.6	335	2	US-08-903-800A-6	Sequence 6, Appl
34	668.5	29.3	340	4	US-09-252-991A-21612	Sequence 21612, A
35	660	29.0	334	4	US-09-674-826B-6	Sequence 6, Appl
36	646	28.4	341	4	US-09-134-001C-3783	Sequence 3783, Ap
37	628	27.6	333	3	US-09-532-803-11	Sequence 11, Appl
38	617	27.1	333	3	US-09-532-803-9	Sequence 9, Appl
39	601.5	26.4	277	4	US-09-548-938A-13	Sequence 13, Appl
40	542.5	23.8	442	4	US-09-540-236-3178	Sequence 3178, A
41	540.5	23.7	350	4	US-09-489-039A-11854	Sequence 11854, A
42	535	23.5	340	4	US-09-328-352-6704	Sequence 6704, Ap
43	532.5	23.4	417	4	US-09-540-236-3470	Sequence 3470, Ap
44	512.5	22.5	359	4	US-09-543-681A-6082	Sequence 6082, Ap
45	504.5	22.1	490	4	US-09-328-352-4616	Sequence 4616, Ap

ALIGNMENTS

RESULT 1

US-09-878-766A-22
; Sequence 22, Application US/09878766A
; Patent No. 660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 22
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple
; OTHER INFORMATION: epitope fusion protein
US-09-878-766A-22

Query Match	100.0%;	Score	2278;	DB	4;	Length	448;
Best Local Similarity	100.0%;	Pred. No.	1.7e-206;				
Matches	448;	Conservative	0;	Mismatches	0;	Indels	0;
Qy	1	MKITGILLALLAVILSACQANYGSMVVKVNGINGFRIGRLAFRRIQNVGVEVTRIN	60				
Db	1	MKITGILLALLAVILSACQANYGSMVVKVNGINGFRIGRLAFRRIQNVGVEVTRIN	60				
Qy	61	DLTDPNMLAHLKYDTTQGRFDGTVEYKEGGFVNGNFIKVSARDPENIDMTDGEIV	120				
Db	61	DLTDPNMLAHLKYDTTQGRFDGTVEYKEGGFVNGNFIKVSARDPENIDMTDGEIV	120				
Qy	121	LEALEGTVEYKGGFVNGNFIKVSARDPENIDMTDGEIVLEIDGTVEYKEGGFVNV	180				
Db	121	LEALEGTVEYKGGFVNGNFIKVSARDPENIDMTDGEIVLEIDGTVEYKEGGFVNV	180				
Qy	181	GQFVKVSAEPEANIDMTDGEIVLEATSFPAKKEAAEKHLHANGAKKVVITAPGNDV	240				
Db	181	GQFVKVSAEPEANIDMTDGEIVLEATSFPAKKEAAEKHLHANGAKKVVITAPGNDV	240				
Qy	241	KTVVFNTHDILDGTETVIGASCTTNCLAPMAKALHDAFGIOKGLMTTHATYTGQML	300				
Db	241	KTVVFNTHDILDGTETVIGASCTTNCLAPMAKALHDAFGIOKGLMTTHATYTGQML	300				
Qy	301	DGPHRGDLRRARAGANIVPNSGTGAATGLVPELNGKLDGAAQRPVPTGVTSLVV	360				

Db 301 DGPFRGGDLRRARAGAANIVPNSTGAAGAIGLVIPELNGKLDGAAQRPVPTGSGVTELVV 360
QY 361 TLDKNVSDVEINAAKAAASNDGFGVTEDPVSSDIVGVSYSGLFDATQTKMEVDGSQLV 420
Db 361 TLDKNVSDVEINAAKAAASNDGFGVTEDPVSSDIVGVSYSGLFDATQTKMEVDGSQLV 420
QY 421 KVVSWYDNEMSYTAQLVRLTEYFAKIAK 448
Db 421 KVVSWYDNEMSYTAQLVRLTEYFAKIAK 448

RESULT 2

US-09-878-766A-12
; Sequence 12, Application US/09878766A
; Patent No. 5660270

GENERAL INFORMATION:

APPLICANT: Potter, Andrew A.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
FILE REFERENCE: 9000-0057
CURRENT APPLICATION NUMBER: US/09/878, 766A
CURRENT FILING DATE: 2001-09-10
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus dysgalactiae
US-09-878-766A-12

Query Match 72.7%; Score 1656.5; DB 4; Length 336;
Best Local Similarity 79.6%; Pred. No. 5e-148;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRIQNVGEVETRLNDLPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRIQNVGEVETRLNDLPNMLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEALEGTVEVKDGGFDVNGKFIKVSAB 147
Db 61 KEGGFVNGNFIKVSARDPE----- 81

QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 82 -----NIDWATDGVIEVLE 95
QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDGTETVIGASCCTN 267
Db 96 ATGFFAKKEAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDGTETVIGASCCTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPFRGGDLRRARAGAANIVPNSTGAA 327
Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPFRGGDLRRARAGAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSDEINAAKAAASNDGFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSDEINAAKAAASNDGFGYTE 275
QY 448 K 448
Db 336 K 336

RESULT 3

US-07-928-462-2
; Sequence 2, Application US/07928462
; Patent No. 5328996

GENERAL INFORMATION:

APPLICANT: Boyle, Michael D.P.
APPLICANT: Lottenberg, Richard
APPLICANT: Broder, Christopher O.
APPLICANT: von Mering, Gregory O.
TITLE OF INVENTION: Bacterial Plasmid Receptors as
TITLE OF INVENTION: Fibrinolytic Agents
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/928,462
FILING DATE: 19920810
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/524,411
FILING DATE: 16-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/330,849
FILING DATE: 29-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF/S&S-13.C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 336 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-928-462-2

Query Match 72.5%; Score 1652.5; DB 1; Length 336;
Best Local Similarity 79.1%; Pred. No. 1.2e-147;
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRIQNVGEVETRLNDLPNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRIQNVGEVETRLNDLPNMLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEALEGTVEVKDGGFDVNGKFIKVSAB 147
Db 61 KEGGFVNGNFIKVSARDPE----- 81
QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 82 -----NIDWATDGVIEVLE 95
QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDGTETVIGASCCTN 267
Db 96 ATGFFAKKEAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDGTETVIGASCCTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPFRGGDLRRARAGAANIVPNSTGAA 327
Db 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPFRGGDLRRARAGAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSDEINAAKAAASNDGFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNVSDEINAAKAAASNDGFGYTE 275

Db 276 DPVSSDIIGMAYGSLFDATQTKVQTVGDNQNLKVVSWYDNEMSYTAQLVRLTYEYFAKIA 335
QY 448 K 448
Db 336 K 336
RESULT 6
US-09-878-766A-14
; Sequence 14, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus iniae
US-09-878-766A-20
Query Match 67.4%; Score 1534.5; DB 4; Length 336;
Best Local Similarity 72.7%; Pred. No. 1.6e-136;
Matches 306; Conservative 17; Mismatches 13; Indels 85; Gaps 1;
QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNNMLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGVETVKGDFVNGKFIKVSAR 147
Db 61 KGGFEVNGS----- 70
QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 71 -----FKVSAEREPANIDWATDGVIEVLE 95
QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGGNDKVTVVNTNHDILDTGTETVISGASCTTN 267
Db 96 ATGFFASKEKAGQIHENGAKKVVITAPGGNDKVTVVNTNHDILDTGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHERGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPHERGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNSVDEINAAKAAASNDSPGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNSVDEINAAKAAASNDSPGYTE 275
QY 388 DPVSSDIVGYSGLFDPATQTKMVEVDSQSLKVVSWYDNEMSYTAQLVRLTYEYFAKIA 447
Db 276 DPVSSDIVGYSGLFDPATQTKMVEVDSQSLKVVSWYDNEMSYTAQLVRLTYEYFAKIA 335
QY 448 K 448
Db 336 K 336
RESULT 8
US-09-878-766A-18
; Sequence 18, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 18
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus parauberis
US-09-878-766A-18
Query Match 67.2%; Score 1530.5; DB 4; Length 336;
Best Local Similarity 73.4%; Pred. No. 3.9e-136;

Db 276 DPVSSDIIGMAYGSLFDATQTKVQTVGDNQNLKVVSWYDNEMSYTAQLVRLTYEYFAKIA 335
QY 448 K 448
Db 336 K 336
RESULT 6
US-09-878-766A-14
; Sequence 14, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-878-766A-14
Query Match 68.4%; Score 1557.5; DB 4; Length 336;
Best Local Similarity 73.9%; Pred. No. 1.1e-138;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNNMLAHLKYDTTQGRFDGTVEV 87
Db 1 MVKVGINGFGRIGRLAFRRIONVEGVETRIINDLTPNNMLAHLKYDTTQGRF----- 54
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGVETVKGDFVNGKFIKVSAR 147
Db 55 ----- 54
QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 55 -----DGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 95
QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGGNDKVTVVNTNHDILDTGTETVISGASCTTN 267
Db 96 ATGFFASKEKAGQIHENGAKKVVITAPGGNDKVTVVNTNHDILDTGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTIHAYTGDQMLDGPHERGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDQMLDGPHERGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNSVDEINAAKAAASNDSPGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLDKNSVDEINAAKAAASNDSPGYTE 275
QY 388 DPVSSDIVGYSGLFDPATQTKMVEVDSQSLKVVSWYDNEMSYTAQLVRLTYEYFAKIA 447
Db 276 DPVSSDIVGYSGLFDPATQTKMVEVDSQSLKVVSWYDNEMSYTAQLVRLTYEYFAKIA 335
QY 448 K 448
Db 336 K 336
RESULT 7
US-09-878-766A-20
; Sequence 20, Application US/09878766A
; Patent No. 6660270
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael

Matches 309; Conservative 13; Mismatches 14; Indels 85; Gaps 2;

QY 28 MVVKVINGFGRIQLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEV 87

Db 1 MVVKVINGFGRIQLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFD----- 55

QY 88 KEGFVNGNFIKVSARDPENIDWATDGEIVLEALSGTVEVKDGGPDVNGKFIKVSAAE 147

Db 56 -----GTVEVKDGGPDVNGKFIKVSAAE 77

QY 148 KDPEQIDWATDGEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGEIVLE 207

Db 78 KDPEQ-----IDWATDGEIVLE 95

QY 208 ATFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVPNTNHDLDGTETVISGASCTTN 267

Db 96 ATGFFAKKAAAEKHLHANGAKKVITAPGGDDVKTVPNTNHDLDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDPHGGDLRRARAGANNIVNSTGAA 327

Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDMILDPHGGDLRRARAGANNIVNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTETLVTLDKNVSVDENAAKAAASNDSPGYTE 387

Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTETLVAVLNKETSVEEINSVMKAAANDSPGYTE 275

QY 388 DPEVSDIVGVSGLFDATQTKMEVDGSQLVKVSWYDNMSYTAQLVTLLEYFAKIA 447

Db 276 DPEVSDIVGVSGLFDATQTKVQVDPGNQLVKVSVDENMSYTAQLVTLLEYFAKIA 335

QY 448 K 448

Db 336 K 336

RESULT 9

US-08-961-083-54

; Sequence 54, Application US/08961083

; Patent No. 6159469

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961,083

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB340P2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 54:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 333 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-961-083-54

Query Match 65.5%; Score 1491; DB 3; Length 333;

Best Local Similarity 72.4%; Pred. NO. 2.1e-132;

Matches 302; Conservative 12; Mismatches 17; Indels 86; Gaps 2;

QY 29 VVKVINGFGRIQLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVEVK 88

Db 1 VVKVINGFGRIQLAFRRIONVEGVETRLNDLTPVMLAHLKYDTTQGRFDGTVEVK 60

QY 89 EGFVNGNFIKVSARDPENIDWATDGEIVLEALSGTVEVKDGGPDVNGKFIKVSAAE 148

Db 61 EGFVNGKFIKVSARDPE----- 80

QY 149 DPEQIDWATDGEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGEIVLEA 208

Db 81 -----QIDWATDGEIVLEA 95

QY 209 TSFFAKKEAAEKHLHANGAKKVITAPGGNDVKTVPNTNHDLDGTETVISGASCTTN 268

Db 96 TGFFAKKEAAEKHL-KGAKKVITAPGGNDVKTVPNTNHDLDGTETVISGASCTTN 154

QY 269 LAPMAKALHDAFGIQKGLMTTHAYTGDMILDPHGGDLRRARAGANNIVNSTGAAK 328

Db 155 LAPMAKALQDNFVGELMTTHAYTGDMILDPHGGDLRRARAGANNIVNSTGAAK 214

QY 329 AIGLVIPELNGKLDGAAQRPVPTGSGVTETLVTLDKNVSVDENAAKAAASNDSPGYTED 388

Db 215 AIGLVIPELNGKLDGSAQRPVPTGSGVTETLVAVLEKNVTDEVNAAKAAASNESGYTED 274

QY 389 PIVSSDIVGVSGLFDATQTKMEVDGSQLVKVSWYDNMSYTAQLVTLLEYFAK 445

Db 275 PIVSSDIVGVSGLFDATQTKVLDVKGOLVKVSWYDNMSYTAQLVTLGILRK 331

RESULT 10

US-09-536-784-54

; Sequence 54, Application US/09536784

; Patent No. 573082

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/536,784

; FILING DATE: 30-Oct-1997

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/961,083

; FILING DATE: OCT-30-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Michelle S. Marks

; REGISTRATION NUMBER: 41,971

; REFERENCE/DOCKET NUMBER: PB340P3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-536-784-54

Query Match 65.5%; Score 1491; DB 4; Length 333;
Best Local Similarity 72.4%; Pred. No. 2.1e-132;
Matches 302; Conservative 12; Mismatches 17; Indels 86; Gaps 2;

29 VVKYINGFGFGRIGLAFRRIONVGEVETRIINDLTDPMHLAKLYDTTQGRDGTVEVK 88
1 VVKYINGFGFGRIGLAFRRIONVGEVETRIINDLTDPMHLAKLYDTTQGRDGTVEVK 60
89 EGGFEVNGFNFKVSAERDPENIDWATDGEIVLEALEGTVEVDKGGFDVNGKFIKVSAX 148
61 EGGFEVNGKFIKVSARDPE-----80
149 DPEQIDWATDGEIVLEIDGTVEVKEGFEVNGQVKVSAEREPANIDWATDGEIVLEA 208
81 -----QIDWATDGEIVLEA 95
209 TSFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVTNTHDILDGTETVISGASCTTNC 268
96 TGFFAKKEAAEKHL-KGSAKKVVITAPGNDVKTVTNTHDILDGTETVISGASCTTNC 154
269 LAPWAKALHDAFGOKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAAK 328
155 LAPWAKALQDNFVGVVEGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAAK 214
329 AIGLVIPELNGKLDGAAQRPVPTGVSVELVTLVKXVSVYDNEMSYTAQLVRLTLEYFAK 445
215 AIGLVIPELNGKLDGSAQRPVPTGVSVELVTLVAVLEKXVTVDEVNAAKAAASNESYGYTED 274
389 PIVSSDIVGVSGLFDATQTKWEVDGSLQVKVSVYDNEMSYTAQLVRLTLEYFAK 445
275 PIVSSDIVGVSGLFDATQTKVLDVGKQLVKVSVYDNEMSYTAQLVRLTLEYFAK 331

RESULT 11
US-09-134-000C-4400
Sequence 4400, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4400
LENGTH: 346
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-4400

Query Match 56.3%; Score 1282; DB 4; Length 346;
Best Local Similarity 62.5%; Pred. No. 1.2e-112;
Matches 262; Conservative 25; Mismatches 46; Indels 86; Gaps 3;

28 MVVKYINGFGFGRIGLAFRRIONVGEVETRIINDLTDPMHLAKLYDTTQGRDGTVEV 87
14 MTVKYINGFGFGRIGLAFRRIONVGEVETRIINDLTDPMHLAKLYDTTQGRDGTVEV 73
88 KEGGFEVNGFNFKVSAERDPENIDWATDGEIVLEALEGTVEVDKGGFDVNGKFIKVSAX 147

74 HEGSENVANGKEIKVLANRNPPELPW-----98
148 XDPEQIDWATDGEIVLEIDGTVEVKEGFEVNGQVKVSAEREPANIDWATDGEIVLE 207
99 -----GEL-----GVDIVLE 108
208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTVTNTHDILDGTETVISGASCTTNC 267
109 CTGFTTSKEAAEKHLTA-GAKRVWISAPGNDVPTIVYTNTHETLTGBETVISGASCTTNC 167
268 CLAPWAKALHDAFGOKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 327
168 CLAPWAKALHDAFGOKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 227
328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLVKXVSVYDNEMSYTAQLVRLTLEYFAK 446
228 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLVKXVSVYDNEMSYTAQLVRLTLEYFAK 287
388 PIVSSDIVGVSGLFDATQTKWEVDGSLQVKVSVYDNEMSYTAQLVRLTLEYFAK 446
288 DEIVSSDIVGVSGLFDATQTKVLDVGKQLVKVSVYDNEMSYTAQLVRLTLEYFAK 346

RESULT 12
US-09-107-532A-4769
Sequence 4769, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4769:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...333
SEQUENCE DESCRIPTION: SEQ ID NO: 4769:
US-09-107-532A-4769

Query Match 55.8%; Score 1270; DB 4; Length 333;
Best Local Similarity 62.1%; Pred. No. 1.5e-111;
Matches 260; Conservative 26; Mismatches 47; Indels 86; Gaps 3;

QY 28 MVVKVINGFGRIGRLAFRIQVGEVTRINDLTDPMHLAKKYDTTQGRFDGTVEV 87
DB 1 MTVKVINGFGRIGRLAFRIQVGEVTRINDLTDPMHLAKKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIVKSAERDPENIDWATDGVVEIVLEAEGTVEVKGDFVNGKFIKVSAAE 147
DB 61 HEGGFVNGNFIVKSAERDPENIDWATDGVVEIVLEAEGTVEVKGDFVNGKFIKVSAAE 85
QY 148 KDPQIDWATDGVVEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207
DB 86 -----GEL-----GVDIVLE 95
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTWVFNTNHDILDTGTVISGASCTTN 267
DB 96 CTGFTTSKAAEKHLTA-GAKRVVISAPGNDVPTVYNTNHTLTGKTVISGASCTTN 154
QY 268 CLAPMAKALHDAPFGIOKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 327
DB 155 CLAPMAKALHDAPFGIOKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 214
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTVTELAVTLDK-NVSVDEINAAKAAASNDGFGYTE 387
DB 215 KAIGLVIPELNGKLDGAAQRPVPTGTVTELAVTLDK-NVSVDEINAAKAAASNDGFGYTE 274
QY 388 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWVDNEMSYTAQLVRLTLEYFAKI 446
DB 275 DEIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWVDNEMSYTAQLVRLTLEYFAKI 333

RESULT 13
US-09-134-001C-5513
; Sequence 5513, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5513
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5513

Query Match 50.0%; Score 1138.5; DB 4; Length 348;
Best Local Similarity 55.7%; Pred. No. 4.2e-99;
Matches 235; Conservative 39; Mismatches 61; Indels 87; Gaps 4;

QY 28 MVVKVINGFGRIGRLAFRIQVGEVTRINDLTDPMHLAKKYDTTQGRFDGTVEV 87
DB 13 MAIKVINGFGRIGRLAFRIQVGEVTRINDLTDPMHLAKKYDTTQGRFDGTVEV 72
QY 88 KEGGFVNGNFIVKSAERDPENIDWATDGVVEIVLEAEGTVEVKGDFVNGKFIKVSAAE 147
DB 73 IEGGFVNGNFIVKSAERDPENIDWATDGVVEIVLEAEGTVEVKGDFVNGKFIKVSAAE 101
QY 148 KDPQIDWATDGVVEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVVEIVLE 207
DB 102 -----IDVLE 107
QY 208 ATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTWVFNTNHDILDTGTVISGASCTTN 267

Query Match 55.8%; Score 1270; DB 4; Length 333;
Best Local Similarity 62.1%; Pred. No. 1.5e-111;
Matches 260; Conservative 26; Mismatches 47; Indels 86; Gaps 3;

QY 28 MVVKVINGFGRIGRLAFRIQVGEVTRINDLTDPMHLAKKYDTTQGRFDGTVEV 86
DB 22 MTVKVINGFGRIGRLAFRIQVGEVTRINDLTDPMHLAKKYDTTQGRFDGTVEV 81
QY 87 KEGGFVNGNFIVKSAERDPENIDWATDGVVEIVLEAEGTVEVKGDFVNGKFIKVSAAE 146
DB 82 ATENGIIVDGEETRVVAEPDASKIPW----- 107
QY 147 EKDPQIDWATDGVVEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVVEIVL 206
DB 108 -----VKE-----NGVDIVL 117
QY 207 EATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTWVFNTNHDILDTGTVISGASCTTN 266
DB 118 ECTGFTTSKAAEKHLTA-GVKEVVISAPAG-AMKIVVNVNDLTDANDKIIISAGSCTT 175
QY 267 NCLAPMAKALHDAPFGIOKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 326
DB 176 NCLAPMAKALHDAPFGIOKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 235
QY 327 KAIGLVIPELNGKLDGAAQRPVPTGTVTELAVTLDK-NVSVDEINAAKAAASNDGFGYTE 384
DB 236 KAIGLVIPELNGKLDGAAQRPVPTGTVTELAVTLDK-NVSVDEINAAKAAASNDGFGYTE 295
QY 385 YTDPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWVDNEMSYTAQLVRLTLEYFAKI 444
DB 296 YDDREIVSGDIIGTTEGSIPTQTEVTTAGDFQVVKVTAVDNEMSYTAQLVRLTLEYFAKI 355

Db 108 CTGFTYDKEKAQAHAIDA-GAKKVLISAPAGKDVKTIVFNTNHDILDTGSETVSGASCTTN 166
QY 268 CLAPMAKALHDAPFGIOKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 327
DB 167 SLAPVAKVLSDFEGLVGEFTTHAYTGDQNTQDAPARKDKRRARAAENIINSTGAA 226
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGTVTELAVTLDK-NVSVDEINAAKAAASNDGFGYTE 386
DB 227 KAIGKVIPEIDGKLDGGAQRPVATGSLTTLTVVLDKQDVTVDQVNSAMKQASDESFGYT 286
QY 387 EDPVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWVDNEMSYTAQLVRLTLEYFAKI 446
DB 297 EDEIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWVDNEMSYTAQLVRLTLEYFAKI 346
QY 447 AK 448
DB 347 SK 348

RESULT 14
US-09-134-000C-4229
; Sequence 4229, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/035,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4229
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4229

Query Match 40.5%; Score 923.5; DB 4; Length 357;
Best Local Similarity 47.2%; Pred. No. 8.5e-79;
Matches 199; Conservative 44; Mismatches 90; Indels 89; Gaps 6;

QY 28 MVVKVINGFGRIGRLAFRIQV-EGVEVTRINDLTDPMHLAKKYDTTQGRFDGTVE 86
DB 22 MTVKVINGFGRIGRLAFRIQV-EGVEVTRINDLTDPMHLAKKYDTTQGRFDGTVE 81
QY 87 KEGGFVNGNFIVKSAERDPENIDWATDGVVEIVLEAEGTVEVKGDFVNGKFIKVSAAE 146
DB 82 ATENGIIVDGEETRVVAEPDASKIPW----- 107
QY 147 EKDPQIDWATDGVVEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVVEIVL 206
DB 108 -----VKE-----NGVDIVL 117
QY 207 EATSFPAKKEAAEKHLHANGAKKVVITAPGNDVKTWVFNTNHDILDTGTVISGASCTTN 266
DB 118 ECTGFTTSKAAEKHLTA-GVKEVVISAPAG-AMKIVVNVNDLTDANDKIIISAGSCTT 175
QY 267 NCLAPMAKALHDAPFGIOKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 326
DB 176 NCLAPMAKALHDAPFGIOKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAAANIVPNSTGAA 235
QY 327 KAIGLVIPELNGKLDGAAQRPVPTGTVTELAVTLDK-NVSVDEINAAKAAASNDGFGYTE 384
DB 236 KAIGLVIPELNGKLDGAAQRPVPTGTVTELAVTLDK-NVSVDEINAAKAAASNDGFGYTE 295
QY 385 YTDPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWVDNEMSYTAQLVRLTLEYFAKI 444
DB 296 YDDREIVSGDIIGTTEGSIPTQTEVTTAGDFQVVKVTAVDNEMSYTAQLVRLTLEYFAKI 355

QY 445 KI 446
Db 356 NL 357

RESULT 15

US-09-634-238-233

; Sequence 233, Application US/09634238

; Patent No. 6544772

; GENERAL INFORMATION:

; APPLICANT: Glenn, Matthew

; APPLICANT: Havukkala, Ilkka J.

; APPLICANT: Bloksberg, Leonard, N.

; APPLICANT: Lubbers, Mark W.

; APPLICANT: Dekker, James

; APPLICANT: Christenson, Anna C.

; APPLICANT: Holland, Rose

; APPLICANT: O'Toole, Paul W.

; APPLICANT: Reid, Julian R.

; APPLICANT: Coolbear, Timothy

; TITLE OF INVENTION: Polynucleotides, materials incorporating

; TITLE OF INVENTION: them and methods for using them.

; FILE REFERENCE: 11000.1043U1

; CURRENT APPLICATION NUMBER: US/09/634,238

; CURRENT FILING DATE: 2000-08-08

; NUMBER OF SEQ ID NOS: 422

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 233

; LENGTH: 340

; TYPE: PRT

; ORGANISM: Lactobacillus rhamnosus

US-09-634-238-233

Query Match 40.3%; Score 918.5; DB 4; Length 340;

Best Local Similarity 46.0%; Pred.No. 2.3e-78;

Matches 196; Conservative 49; Mismatches 88; Indels 93; Gaps 7;

QY 28 MVKVGINGCRIGLAFRIQV-----EGVEVTRINDLPNLAHLKYDTTQGRFDG 83

Db 1 MTVKIGINGCRIGLAFRIYEIGAKSNDIQVAINDLTSPMLAHLKYDSTHGTFFG 60

QY 84 TVEVKEGGFEVNGNFIKVSARDPDNDWATDGVIEVLEALEGTVEVKGDFVNGKFIK 143

Db 61 EVSATDNGIIVDGEYVYAEPAQNIPIW-----VKN----- 92

QY 144 VSAEKDPEQIDWATDGVIEVLEIDGTVEVKEGGFEVNGQFVKVSAAREPANIDWATDGE 203

Db 93 -----DGVD 96

QY 204 IVLEATSPFAKKEAAEKHLHANGAKVITAPGNDVKTIVVFNTHDILDGTETVIGAS 263

Db 97 YVLECTGYTSAEKSQAHLDA-GAKRVLISAPACK-IKTIVYVNDTTLNADDKIVSAGS 154

QY 264 CTTCNCLAPMAKALHDARGIQKGLMTTHAYTGDMDLDPHGGDLRRARAGANIVPNS 323

Db 155 CTTCNCLAPMAYFLNQEPGIEVGTWTTVHAYTSTQMLDGPVGGNLRARSAANTIPHS 214

QY 324 TGAAKAIGLIVPELNGKLDGAAQRPVPTGSVTTELVTTL-DKNVSVDEINAAMK--AASN 380

Db 215 TGAAKAIGLIVPELNGKLDGAAQRPVPTGSVTTELVTTL-DKNVSVDEINAAMK--AASN 380

QY 381 DSFGYTEPIVSSDIIVGVSGLSDATQTKVNEVDGSQLVKVSWYDNEMSYTAQLVRTL 440

Db 275 PSFGWNEDEIVSSDVIGTGYTSIEDPTQTEVTAGDYQLVKTVAWYDNEYGFTQCMIRTL 334

QY 441 EYFAKI 446

Db 335 LKFATL 340

Search completed: March 29, 2004, 16:36:43
Job time : 33.5224 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 16:35:22 ; Search time 73.5522 Seconds
(without alignments)
1593.758 Million cell updates/sec

Title: US-10-650-369-22

Perfect score: 2278

Sequence: 1 MKKITGIALLLAIVLSAC.....EMSYTAQLVRLTEYFAKIAK 448

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PTCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PTCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2278	100.0	448	9	US-09-878-766A-22
2	1656.5	72.7	336	9	US-09-878-766A-12
3	1656.5	72.7	336	10	US-09-878-781-12
4	1656.5	72.7	336	14	US-10-134-297-4
5	1655.5	72.7	336	12	US-10-282-122A-74379
6	1650	72.4	335	10	US-09-878-781-14
7	1562.5	68.6	336	10	US-09-878-781-16
8	1559.5	68.5	336	9	US-09-878-766A-16
9	1559.5	68.5	336	10	US-09-878-781-8
10	1559.5	68.5	336	14	US-10-134-297-8
11	1557.5	68.4	336	9	US-09-878-766A-14
12	1557.5	68.4	336	10	US-09-878-781-6
13	1557.5	68.4	336	9	US-09-878-781-6
14	1535	67.4	359	9	US-09-878-781-6
15	1535	67.4	359	12	US-10-282-122A-74186

16	1534.5	67.4	336	9	US-09-878-766A-20
17	1534.5	67.4	336	10	US-09-878-781-12
18	1534.5	67.4	336	14	US-10-134-297-12
19	1534	67.3	359	9	US-09-815-242-13593
20	1530.5	67.2	336	9	US-09-878-766A-18
21	1530.5	67.2	336	10	US-09-878-781-10
22	1530.5	67.2	336	14	US-10-134-297-10
23	1521	66.8	337	12	US-10-282-122A-72254
24	1491	65.8	333	9	US-09-765-272-54
25	1357.5	59.6	336	15	US-10-369-493-18524
26	1282	56.3	333	9	US-09-815-242-10847
27	1282	56.3	333	12	US-10-282-122A-57294
28	1280	56.2	337	15	US-10-369-493-18342
29	1244.5	53.8	335	12	US-10-282-122A-53266
30	1205.5	52.9	336	12	US-10-282-122A-60492
31	1203	52.8	334	12	US-10-282-122A-51519
32	1166	51.2	335	12	US-10-282-122A-71734
33	1145	50.3	334	12	US-10-282-122A-65661
34	1141	50.1	357	12	US-10-282-122A-65001
35	1138.5	50.0	336	12	US-10-282-122A-70890
36	1132.5	49.7	336	9	US-09-815-242-5618
37	1132.5	49.7	336	9	US-09-815-242-12540
38	1132.5	49.7	336	12	US-10-282-122A-43997
39	941.5	41.3	333	12	US-10-282-122A-56828
40	918.5	40.3	340	15	US-10-264-213-141
41	916.5	40.2	340	9	US-09-971-536-65
42	916.5	40.2	340	15	US-10-264-213-252
43	908.5	39.9	332	12	US-10-282-122A-60108
44	900.5	39.5	336	12	US-10-282-122A-57807
45	870.5	38.2	337	12	US-10-282-122A-64191

ALIGNMENTS

RESULT 1
US-09-878-766A-22
; Sequence 22, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878, 766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GapC multiple
; OTHER INFORMATION: epitope fusion protein
US-09-878-766A-22

Query Match	100.0%	Score	2278	DB	9	Length	448
Best Local Similarity	100.0%	Pred. No.	3.4e-199				
Matches	448	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MKKITGIALLLAIVLSACQANYGSMVVKVINGFGRIQVGEVETRN	60				
Db	1	MKKITGIALLLAIVLSACQANYGSMVVKVINGFGRIQVGEVETRN	60				
Qy	61	DLTDPNMLAHLKYDTQGRFGTVEVKEGSGVNGNFKVSAERDPENIDATGVEIV	120				
Db	61	DLTDPNMLAHLKYDTQGRFGTVEVKEGSGVNGNFKVSAERDPENIDATGVEIV	120				
Qy	121	LEALEGTVEKDGDFVNGKFKVSAEKDPEQIDWATGVEIVLEIDGTVEVKEGSGFVN	180				

Db 121 LEALGTEVVKGGFDVNGKFIKVSAREKQIDWATDGVIEVLIDGTVEVKEGFEVN 180
 QY 181 GQFVKVSAEREPANIDWATDGVIEVLIDGTVEVKEGFEVN 240
 Db 181 GQFVKVSAEREPANIDWATDGVIEVLIDGTVEVKEGFEVN 240
 QY 241 KTVVFNTHDILDTGTETVSGASCTNCLAPMAKALHDAFGIQKGLMTTHIAYTGDQML 300
 Db 241 KTVVFNTHDILDTGTETVSGASCTNCLAPMAKALHDAFGIQKGLMTTHIAYTGDQML 300
 QY 301 DGHRRGDLRRARAGANIVPNSGTGAAGATGLVPELNGKLDGAAQRPVPTGSGVTVELV 360
 Db 301 DGHRRGDLRRARAGANIVPNSGTGAAGATGLVPELNGKLDGAAQRPVPTGSGVTVELV 360
 QY 361 TLDKNVSVDEINAAKASNDSPGYTETDPTVSSDIVSGVSLFDATQTKMVEVDGSQLV 420
 Db 361 TLDKNVSVDEINAAKASNDSPGYTETDPTVSSDIVSGVSLFDATQTKMVEVDGSQLV 420
 QY 421 KTVSWYDNEMSYTAQLVRLTLEYFAKIAK 448
 Db 421 KTVSWYDNEMSYTAQLVRLTLEYFAKIAK 448

RESULT 2
 US-09-878-766A-12
 ; Sequence 12, Application US/09878766A
 ; Patent No. US2002004928A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Potter, Andrew A.
 ; APPLICANT: Perez-Casal, Jose
 ; APPLICANT: Fontaine, Michael
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
 ; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
 ; FILE REFERENCE: 9000-0057
 ; CURRENT APPLICATION NUMBER: US/09/878,766A
 ; CURRENT FILING DATE: 2001-09-10
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 12
 ; LENGTH: 336
 ; TYPE: PRT
 ; ORGANISM: Streptococcus dysgalactiae
 US-09-878-766A-12

Query Match 72.7%; Score 1656.5; DB 9; Length 336;
 Best Local Similarity 79.6%; Pred. No. 1.2e-142;
 Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDNMLAHLKYDTTQGRFDGTVEV 87
 Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDNMLAHLKYDTTQGRFDGTVEV 60
 QY 88 KEGGFEVNGNFIKVSAREDPENIDWATDGVIEVLEALGTEVVKDGGFDVNGKFIKVSARE 147
 Db 61 KEGGFEVNGNFIKVSAREDPENIDWATDGVIEVLEALGTEVVKDGGFDVNGKFIKVSARE 81
 QY 148 KDEQIDWATDGVIEVLIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
 Db 82 -----NIDWATDGVIEVLE 95
 QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGGNDVKTWVNTNHDILDTGTETVSGASCTTN 267
 Db 96 ATGFFAKKEAEKHLHANGAKKVVITAPGGNDVKTWVNTNHDILDTGTETVSGASCTTN 155
 QY 268 CLAPMAKALHDAFGIQKGLMTTHIAYTGDQMLDGPHRGGDLRRARAGANIVPNSGTAA 327
 Db 156 CLAPMAKALHDAFGIQKGLMTTHIAYTGDQMLDGPHRGGDLRRARAGANIVPNSGTAA 215
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTVELVTLTKNVSVDEINAAKASNDSPGYTE 387
 Db 61 KEGGFEVNGNFIKVSAREDPENIDWATDGVIEVLEALGTEVVKDGGFDVNGKFIKVSARE 81
 QY 148 KDEQIDWATDGVIEVLIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
 Db 82 -----NIDWATDGVIEVLE 95
 QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGGNDVKTWVNTNHDILDTGTETVSGASCTTN 267
 Db 96 ATGFFAKKEAEKHLHANGAKKVVITAPGGNDVKTWVNTNHDILDTGTETVSGASCTTN 155
 QY 268 CLAPMAKALHDAFGIQKGLMTTHIAYTGDQMLDGPHRGGDLRRARAGANIVPNSGTAA 327
 Db 156 CLAPMAKALHDAFGIQKGLMTTHIAYTGDQMLDGPHRGGDLRRARAGANIVPNSGTAA 215
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTVELVTLTKNVSVDEINAAKASNDSPGYTE 387
 Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTVELVTLTKNVSVDEINAAKASNDSPGYTE 275
 QY 388 DPVSSDIVSGVSLFDATQTKMVEVDGSQLVTKVSWYDNEMSYTAQLVRLTLEYFAKIA 447

Db 276 DPVSSDIVSGVSLFDATQTKMVEVDGSQLVTKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
 QY 448 K 448
 Db 336 K 336

RESULT 3
 US-09-878-781-4
 ; Sequence 4, Application US/09878781
 ; Publication No. US20030082781A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolton, Alexandra J.
 ; APPLICANT: Perez-Casal, Jose
 ; APPLICANT: Fontaine, Michael
 ; APPLICANT: Potter, Andrew A.
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
 ; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
 ; FILE REFERENCE: 9000-0055
 ; CURRENT APPLICATION NUMBER: US/09/878,781
 ; CURRENT FILING DATE: 2002-09-10
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 336
 ; TYPE: PRT
 ; ORGANISM: Streptococcus dysgalactiae
 US-09-878-781-4

Query Match 72.7%; Score 1656.5; DB 10; Length 336;
 Best Local Similarity 79.6%; Pred. No. 1.2e-142;
 Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDNMLAHLKYDTTQGRFDGTVEV 87
 Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRINDLTDNMLAHLKYDTTQGRFDGTVEV 60
 QY 88 KEGGFEVNGNFIKVSAREDPENIDWATDGVIEVLEALGTEVVKDGGFDVNGKFIKVSARE 147
 Db 61 KEGGFEVNGNFIKVSAREDPENIDWATDGVIEVLEALGTEVVKDGGFDVNGKFIKVSARE 81
 QY 148 KDEQIDWATDGVIEVLIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
 Db 82 -----NIDWATDGVIEVLE 95
 QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGGNDVKTWVNTNHDILDTGTETVSGASCTTN 267
 Db 96 ATGFFAKKEAEKHLHANGAKKVVITAPGGNDVKTWVNTNHDILDTGTETVSGASCTTN 155
 QY 268 CLAPMAKALHDAFGIQKGLMTTHIAYTGDQMLDGPHRGGDLRRARAGANIVPNSGTAA 327
 Db 156 CLAPMAKALHDAFGIQKGLMTTHIAYTGDQMLDGPHRGGDLRRARAGANIVPNSGTAA 215
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTVELVTLTKNVSVDEINAAKASNDSPGYTE 387
 Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTVELVTLTKNVSVDEINAAKASNDSPGYTE 275
 QY 388 DPVSSDIVSGVSLFDATQTKMVEVDGSQLVTKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
 Db 276 DPVSSDIVSGVSLFDATQTKMVEVDGSQLVTKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
 QY 448 K 448
 Db 336 K 336

RESULT 4
 US-10-134-297-4
 ; Sequence 4, Application US/10134297
 ; Publication No. US20030165524A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolton, Alexandra J.

; APPLICANT: Perez-Casal, Jose
 ; APPLICANT: Fontaine, Michael
 ; APPLICANT: Potter, Andrew A.
 ; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
 ; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
 ; FILE REFERENCE: 9000-0055.20 US/10/134,297
 ; CURRENT APPLICATION NUMBER: US/10/134,297
 ; CURRENT FILING DATE: 2002-04-26
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 336
 ; TYPE: PRT
 ; ORGANISM: Streptococcus dysgalactiae
 US-10-134-297-4

Query Match 72.7%; Score 1656.5; DB 14; Length 336;
 Best Local Similarity 79.6%; Pred. No. 1.2e-142;
 Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;
 QY 28 MVKVGINGFGRIGRLAFRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRPDGTVEV 87
 DB 1 MVKVGINGFGRIGRLAFRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRPDGTVEV 60
 QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKGDFVNGKFIKVS 147
 DB 61 KEGGFEVNGNFIKVSARDPE----- 81
 QY 148 KQPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEV 207
 DB 82 -----NIDWATDGVIEVLE 95
 QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDGTETVSGASCTTN 267
 DB 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDGTETVSGASCTTN 155
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 327
 DB 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 215
 QY 328 KAIGLVPELNGKLDGAARQVPVPTGVTGLVTLVDKNVSDVEINAAKKAANDSGFYTE 387
 DB 216 KAIGLVPELNGKLDGAARQVPVPTGVTGLVTLVDKNVSDVEINAAKKAANDSGFYTE 275
 QY 388 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447
 DB 276 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335
 QY 448 K 448
 DB 336 K 336

RESULT 5
 US-10-282-122A-74379
 ; Sequence 74379, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 74379
 ; LENGTH: 336
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pyogenes
 US-10-282-122A-74379

Query Match 72.7%; Score 1655.5; DB 12; Length 336;
 Best Local Similarity 79.3%; Pred. No. 1.5e-142;
 Matches 334; Conservative 1; Mismatches 1; Indels 85; Gaps 1;
 QY 28 MVKVGINGFGRIGRLAFRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRPDGTVEV 87
 DB 1 MVKVGINGFGRIGRLAFRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRPDGTVEV 60
 QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEVKGDFVNGKFIKVS 147
 DB 61 KEGGFEVNGNFIKVSARDPE----- 81
 QY 148 KQPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEV 207
 DB 82 -----NIDWATDGVIEVLE 95
 QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDGTETVSGASCTTN 267
 DB 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDGTETVSGASCTTN 155
 QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 327
 DB 156 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 215
 QY 328 KAIGLVPELNGKLDGAARQVPVPTGVTGLVTLVDKNVSDVEINAAKKAANDSGFYTE 387
 DB 216 KAIGLVPELNGKLDGAARQVPVPTGVTGLVTLVDKNVSDVEINAAKKAANDSGFYTE 275
 QY 388 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 447
 DB 276 DPIVSSDIVGVSGLFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRTLEYFAKIA 335
 QY 448 K 448
 DB 336 K 336

RESULT 6
 US-09-878-781-14
 ; Sequence 14, Application US/09878781
 ; Publication No. US20030082781A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolton, Alexandra J.

APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
APPLICANT: Potter, Andrew A.
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
TITLE OF INVENTION: STREPTOCOCCUS INFECTION
FILE REFERENCE: 9000-0055
CURRENT FILING DATE: 2002-09-10
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 14
LENGTH: 335
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SpyGapC
OTHER INFORMATION: protein
US-09-878-781-14

Query Match 72.4%; Score 1650; DB 10; Length 335;
Best Local Similarity 79.1%; Pred. No. 4.9e-142;
Matches 333; Conservative 2; Mismatches 0; Indels 86; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 87
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDMATDGVIEVLEALEGTVKDGDFVNGKFIKVSAB 147
DB 61 KEGGFVNGNFIKVSARDPE----- 81
QY 148 KDPQIDMATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDMATDGVIEVLE 207
DB 82 -----IDMATDGVIEVLE 94
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDGTETVISGASCTTN 267
DB 95 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDGTETVISGASCTTN 154
QY 268 CLAPMAKALHDAGIQKGLMTTHAYTGDQMIIDGPHRGGLRRARAGAAIVPNSTGAA 327
DB 155 CLAPMAKALHDAGIQKGLMTTHAYTGDQMIIDGPHRGGLRRARAGAAIVPNSTGAA 214
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDSEINAAMKAASNDSFGYTE 387
DB 215 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDSEINAAMKAASNDSFGYTE 274
QY 388 DPIVSSDIVGVSGLSFDTATQTKMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
DB 275 DPIVSSDIVGVSGLSFDTATQTKMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 334
QY 448 K 448
DB 335 K 335

RESULT 7
US-09-878-781-16
Sequence 16, Application US/09878781
Publication No. US20030082781A1
GENERAL INFORMATION:
APPLICANT: Bolton, Alexandra J.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
APPLICANT: Potter, Andrew A.
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
FILE REFERENCE: 9000-0055
CURRENT FILING DATE: 2002-09-10
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 16

APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
APPLICANT: Potter, Andrew A.
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
TITLE OF INVENTION: STREPTOCOCCUS INFECTION
FILE REFERENCE: 9000-0055
CURRENT FILING DATE: 2002-09-10
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 14
LENGTH: 335
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SpyGapC
OTHER INFORMATION: protein
US-09-878-781-14

Query Match 72.4%; Score 1650; DB 10; Length 335;
Best Local Similarity 79.1%; Pred. No. 4.9e-142;
Matches 333; Conservative 2; Mismatches 0; Indels 86; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 87
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDMATDGVIEVLEALEGTVKDGDFVNGKFIKVSAB 147
DB 61 KEGGFVNGNFIKVSARDPE----- 81
QY 148 KDPQIDMATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDMATDGVIEVLE 207
DB 82 -----IDMATDGVIEVLE 94
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDGTETVISGASCTTN 267
DB 95 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDGTETVISGASCTTN 154
QY 268 CLAPMAKALHDAGIQKGLMTTHAYTGDQMIIDGPHRGGLRRARAGAAIVPNSTGAA 327
DB 155 CLAPMAKALHDAGIQKGLMTTHAYTGDQMIIDGPHRGGLRRARAGAAIVPNSTGAA 214
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDSEINAAMKAASNDSFGYTE 387
DB 215 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDSEINAAMKAASNDSFGYTE 274
QY 388 DPIVSSDIVGVSGLSFDTATQTKMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
DB 275 DPIVSSDIVGVSGLSFDTATQTKMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 334
QY 448 K 448
DB 335 K 335

RESULT 7
US-09-878-781-16
Sequence 16, Application US/09878781
Publication No. US20030082781A1
GENERAL INFORMATION:
APPLICANT: Bolton, Alexandra J.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
APPLICANT: Potter, Andrew A.
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GAPC PROTEIN AGAINST
FILE REFERENCE: 9000-0055
CURRENT FILING DATE: 2002-09-10
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 16

LENGTH: 336
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SeqGapC
OTHER INFORMATION: protein
US-09-878-781-16

Query Match 68.6%; Score 1562.5; DB 10; Length 336;
Best Local Similarity 76.2%; Pred. No. 4.8e-134;
Matches 321; Conservative 2; Mismatches 13; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 87
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDMATDGVIEVLEALEGTVKDGDFVNGKFIKVSAB 147
DB 61 KEGGFVNGNFIKVSARDPE----- 81
QY 148 KDPQIDMATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDMATDGVIEVLE 207
DB 82 -----NIDMATDGVIEVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDGTETVISGASCTTN 267
DB 96 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAGIQKGLMTTHAYTGDQMIIDGPHRGGLRRARAGAAIVPNSTGAA 327
DB 156 CLAPMAKALHDAGIQKGLMTTHAYTGDQMIIDGPHRGGLRRARAGAAIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDSEINAAMKAASNDSFGYTE 387
DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLTKNVSVDSEINAAMKAASNDSFGYTE 275
QY 388 DPIVSSDIVGVSGLSFDTATQTKMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
DB 276 DPIVSSDIVGVSGLSFDTATQTKMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
DB 336 K 336

RESULT 8
US-09-878-766A-16
Sequence 16, Application US/09878766A
Patent No. US20020044928A1
GENERAL INFORMATION:
APPLICANT: Potter, Andrew A.
APPLICANT: Perez-Casal, Jose
APPLICANT: Fontaine, Michael
TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
FILE REFERENCE: 9000-0057
CURRENT FILING DATE: 2001-09-10
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 16
LENGTH: 336
TYPE: PRT
ORGANISM: Streptococcus uberis
US-09-878-766A-16

Query Match 68.5%; Score 1559.5; DB 9; Length 336;
Best Local Similarity 73.9%; Pred. No. 9e-134;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 87
DB 1 MVVKVINGFGRIGRLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTVEV 60

QY 88 KEGGFVNGNFIKVSAREDPENIDWATDGVBEIVLEALLEGTVKDGDFVNGKFIKVSAB 147
Db 61 KOGGFVNGNFIKVSAREDPENIDWATDGVBEIVLEALLEGTVKDGDFVNGKFIKVSAB 81
QY 148 KDPQIDWATDGVBEIVLEALLEGTVKDGDFVNGKFIKVSAB 207
Db 82 -----NIDWATDGVBEIVLE 95
QY 208 ATSPFAKKAEEKHLHANGAKKVVITAPGGNDVKTVPNTNHDLDGTETVISGASCTTN 267
Db 96 ATGFFAKKAEEKHLHANGAKKVVITAPGGDDVKTVPNTNHDLDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDFAFGIOKGLMTTHAYTGDQMLDGHRRGGLRRARAGASNIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKGLMTTHAYTGDQMLDGHRRGGLRRARAGASNIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVDKNVSVDEINAAKAAASNDSEFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVAVLEKETSVEEINAAKAAASNDSEFGYTE 275
QY 388 DPVSSDIVGVSGLSFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPVSSDIIGMAYGSLFDATQTKVQVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 9
US-09-878-781-8
; Sequence 8, Application US/098781
; Publication No. US2003082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781
; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 8
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-09-878-781-8

Query Match 68.5%; Score 1559.5; DB 10; Length 336;
Best Local Similarity 73.9%; Pred. No. 9e-134;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIGRLAFRIQNVGEVTRINDLTPNMLAHLLKYDTTQGRFDGTVEV 87
Db 1 MVKVGINGFGRIGRLAFRIQNVGEVTRINDLTPNMLAHLLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSAREDPENIDWATDGVBEIVLEALLEGTVKDGDFVNGKFIKVSAB 147
Db 61 KOGGFVNGNFIKVSAREDPENIDWATDGVBEIVLEALLEGTVKDGDFVNGKFIKVSAB 81
QY 148 KDPQIDWATDGVBEIVLEALLEGTVKDGDFVNGKFIKVSAB 207
Db 82 -----NIDWATDGVBEIVLE 95
QY 208 ATSPFAKKAEEKHLHANGAKKVVITAPGGNDVKTVPNTNHDLDGTETVISGASCTTN 267
Db 96 ATGFFAKKAEEKHLHANGAKKVVITAPGGDDVKTVPNTNHDLDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDFAFGIOKGLMTTHAYTGDQMLDGHRRGGLRRARAGASNIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKGLMTTHAYTGDQMLDGHRRGGLRRARAGASNIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVDKNVSVDEINAAKAAASNDSEFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVAVLEKETSVEEINAAKAAASNDSEFGYTE 275
QY 388 DPVSSDIVGVSGLSFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPVSSDIIGMAYGSLFDATQTKVQVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

Db 156 CLAPMAKALQDNFVGKGLMTTHAYTGDQMLDGHRRGGLRRARAGASNIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVDKNVSVDEINAAKAAASNDSEFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVAVLEKETSVEEINAAKAAASNDSEFGYTE 275
QY 388 DPVSSDIVGVSGLSFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPVSSDIIGMAYGSLFDATQTKVQVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 10
US-10-134-297-8
; Sequence 8, Application US/10134297
; Publication No. US2003016524A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; FILE REFERENCE: 9000-0055.20
; CURRENT APPLICATION NUMBER: US/10/134,297
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 8
; TYPE: PRT
; ORGANISM: Streptococcus uberis
US-10-134-297-8

Query Match 68.5%; Score 1559.5; DB 14; Length 336;
Best Local Similarity 73.9%; Pred. No. 9e-134;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIGRLAFRIQNVGEVTRINDLTPNMLAHLLKYDTTQGRFDGTVEV 87
Db 1 MVKVGINGFGRIGRLAFRIQNVGEVTRINDLTPNMLAHLLKYDTTQGRFDGTVEV 60
QY 88 KEGGFVNGNFIKVSAREDPENIDWATDGVBEIVLEALLEGTVKDGDFVNGKFIKVSAB 147
Db 61 KOGGFVNGNFIKVSAREDPENIDWATDGVBEIVLEALLEGTVKDGDFVNGKFIKVSAB 81
QY 148 KDPQIDWATDGVBEIVLEALLEGTVKDGDFVNGKFIKVSAB 207
Db 82 -----NIDWATDGVBEIVLE 95
QY 208 ATSPFAKKAEEKHLHANGAKKVVITAPGGNDVKTVPNTNHDLDGTETVISGASCTTN 267
Db 96 ATGFFAKKAEEKHLHANGAKKVVITAPGGDDVKTVPNTNHDLDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDFAFGIOKGLMTTHAYTGDQMLDGHRRGGLRRARAGASNIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKGLMTTHAYTGDQMLDGHRRGGLRRARAGASNIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVDKNVSVDEINAAKAAASNDSEFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVAVLEKETSVEEINAAKAAASNDSEFGYTE 275
QY 388 DPVSSDIVGVSGLSFDATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
Db 276 DPVSSDIIGMAYGSLFDATQTKVQVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 11

US-09-878-766A-14
; Sequence 14, Application US/09878766A
; Patent No. US20020044928A1
; GENERAL INFORMATION:
; APPLICANT: Potter, Andrew A.
; APPLICANT: Fontaine, Michael
; APPLICANT: Perez-Casal, Jose
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH CHIMERIC GAPC PROTEIN
; TITLE OF INVENTION: AGAINST STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0057
; CURRENT APPLICATION NUMBER: US/09/878,766A
; CURRENT FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-878-766A-14

Query Match 68.4%; Score 1557.5; DB 9; Length 336;
Best Local Similarity 73.9%; Pred. No. 1.4e-133;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFPGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRF----- 54
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEALEGTVKGGDFVNGKFIKVS 147
Db 55 ----- 54
QY 148 KDPQIDWATDGVETVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVETV 207
Db 55 -----DGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVETV 95
QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDGTETVIGSCTTN 267
Db 96 ATGFASKERKAGQIHENGAKKVVITAPGNDVKTVVFNTHDILDGTETVIGSCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTETLWTLDKNVSVDENAAKAAASNDSPGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTETLWTLDKNVSVDENAAKAAASNDSPGYTE 275
QY 388 DPIVSSDIVGSGSLFDATQTKWVEDGSQLVKVSWYDNEMSYTAQLVRLTLEYPAKIA 447
Db 276 DPIVSSDIVGSGSLFDATQTKVQTVGNGQLVKVSWYDNEMSYTSQVRLTLEYPAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 12

US-09-878-781-6
; Sequence 6, Application US/09878781
; Publication No. US20030082781A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055
; CURRENT APPLICATION NUMBER: US/09/878,781

; CURRENT FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-878-781-6
Query Match 68.4%; Score 1557.5; DB 10; Length 336;
Best Local Similarity 73.9%; Pred. No. 1.4e-133;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFPGTVEV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRF----- 54
QY 88 KEGGFEVNGNFIKVSARDPENIDWATDGVETVLEALEGTVKGGDFVNGKFIKVS 147
Db 55 ----- 54
QY 148 KDPQIDWATDGVETVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVETV 207
Db 55 -----DGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVETV 95
QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDGTETVIGSCTTN 267
Db 96 ATGFASKERKAGQIHENGAKKVVITAPGNDVKTVVFNTHDILDGTETVIGSCTTN 155
QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTETLWTLDKNVSVDENAAKAAASNDSPGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTETLWTLDKNVSVDENAAKAAASNDSPGYTE 275
QY 388 DPIVSSDIVGSGSLFDATQTKWVEDGSQLVKVSWYDNEMSYTAQLVRLTLEYPAKIA 447
Db 276 DPIVSSDIVGSGSLFDATQTKVQTVGNGQLVKVSWYDNEMSYTSQVRLTLEYPAKIA 335
QY 448 K 448
Db 336 K 336

RESULT 13

US-10-134-297-6
; Sequence 6, Application US/10134297
; Publication No. US20030165524A1
; GENERAL INFORMATION:
; APPLICANT: Bolton, Alexandra J.
; APPLICANT: Perez-Casal, Jose
; APPLICANT: Fontaine, Michael
; APPLICANT: Potter, Andrew A.
; TITLE OF INVENTION: IMMUNIZATION OF DAIRY CATTLE WITH GapC PROTEIN AGAINST
; TITLE OF INVENTION: STREPTOCOCCUS INFECTION
; FILE REFERENCE: 9000-0055.20
; CURRENT APPLICATION NUMBER: US/10/134,297
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-10-134-297-6

Query Match 68.4%; Score 1557.5; DB 14; Length 336;
Best Local Similarity 73.9%; Pred. No. 1.4e-133;
Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;
QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFPGTVEV 87

Db 1 MVKVGINGRIGRLAFRIQNVGVEVTRINDLTPNLAHLKYDTTQGRF 54
Qy 88 KGGFEVNGNFIKVSAREDPENIDWATDGVIEVLEALGTVKDGDFVNGKFIKVSAB 147
Db 55 ----- 54
Qy 148 KQPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 55 -----DGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 95
Qy 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKVTVFNTNHDILDTGTTETVLSGASCTTN 267
Db 96 ATGFASKKAGQHENGAKKVVITAPGNDVKVTVFNTNHDILDTGTTETVLSGASCTTN 155
Qy 268 CLAPMAKALHDFAFGIQKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAANIVPNSGTAA 327
Db 156 CLAPMAKALQDNFVGKQGLMTTHAYTGDQMLDGPGRGGDLRRARAGAANIVPNSGTAA 215
Qy 328 KAIGLVIPELNGKLDGAAQORVPVPTGTSVTELVTLDKVSVDEINAAKKAASNDSPGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQORVPVPTGTSVTELVTLATLEKDVTVBEVNAAKKAASNDSPGYTE 275
Qy 388 DPVSSDIVGVSGLSFDATQTKMVEVDGSQLVKVSWYDNEMSYTAQLVRLTYEYFAKIA 447
Db 276 DPVSSDIVGVSGLSFDATQTKVQVVDGNQLVKVSWYDNEMSYTSQVRLTYEYFAKIA 335
Qy 448 K 448
Db 336 K 336

RESULT 14

US-09-815-242-13169
; Sequence 13169, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13169
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13169

Query Match

67.4%; Score 1535; DB 9; Length 359;

Best Local Similarity 73.6%; Pred. No. 1.7e-131;
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;
Qy 28 MVKVGINGRIGRLAFRIQNVGVEVTRINDLTPNLAHLKYDTTQGRPDGTVGV 87
Db 25 MVKVGINGRIGRLAFRIQNVGVEVTRINDLTPVLAHLKYDTTQGRPDGTVGV 84
Qy 88 KGGFEVNGNFIKVSAREDPENIDWATDGVIEVLEALGTVKDGDFVNGKFIKVSAB 147
Db 85 KGGFEVNGKFIKVSAREDP----- 105
Qy 148 KQPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 106 -----QIDWATDGVIEVLE 119
Qy 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKVTVFNTNHDILDTGTTETVLSGASCTTN 267
Db 120 ATGFASKKAEKHL-KGGAKKVVITAPGNDVKVTVFNTNHDVLDGTETVLSGASCTTN 178
Qy 268 CLAPMAKALHDFAFGIQKGLMTTHAYTGDQMLDGPGRGGDLRRARAGAANIVPNSGTAA 327
Db 179 CLAPMAKALQDNFVGVEGLMTTHAYTGDQMLDGPGRGGDLRRARAGAANIVPNSGTAA 238
Qy 328 KAIGLVIPELNGKLDGAAQORVPVPTGTSVTELVTLDKVSVDEINAAKKAASNDSPGYTE 387
Db 239 KAIGLVIPELNGKLDGSAQORVPVPTGTSVTELVAVLEKNVTVDEVNAAKKAASNDSPGYTE 298
Qy 388 DPVSSDIVGVSGLSFDATQTKMVEVDGSQLVKVSWYDNEMSYTAQLVRLTYEYFAKIA 447
Db 299 DPVSSDIVGVSGLSFDATQTKVLDVDRKQLVKVSWYDNEMSYTAQLVRLTYEYFAKIA 358
Qy 448 K 448
Db 359 K 359

RESULT 15

US-10-282-122A-74186
; Sequence 74186, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22


```
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 74186
/ LENGTH: 359
/ TYPE: PRT
/ ORGANISM: Streptococcus pneumoniae
/ US-10-282-122A-74186

Query Match      67.4%; Score 1595; DB 12; Length 359;
Best Local Similarity 73.6%; Pred. No. 1.7e-131;
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

Qy 28 MVVKVGVNGRIGRLAFRIQNVGVETRIINDLDPNLAHLKLYDTTQGRFDGTVEV 87
Db 25 MVVKVGVNGRIGRLAFRIQNVGVETRIINDLDPNLAHLKLYDTTQGRFDGTVEV 84
Qy 88 KEGGFEVNGNFIKVSARDPENIDWATDGVVEIVLEAEGTVEVKGDFVNGKFIKVS 147
Db 85 KEGGFEVNGKFIKVSARDPE----- 105
Qy 148 KDFEIQDWDATDGVVEIVLEIDGTVEVKEGFEVNGQFVKVSAEREPAIDWATDGV 207
Db 106 -----QIDWATDGVVEIVLE 119
Qy 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGGNDVKTVVFNTHDILDGTETVISGASCTTN 267
Db 120 ATGFFAKKEAAEKHL-KGGAKKVVITAPGGNDVKTVVFNTHDVLDTETVISGASCTTN 178
Qy 268 CLAPMAKALHDAPGIQKGLMTTHAYTGDQMIIDGPHRGGLRRAGAGANIVPNSGTAA 327
Db 179 CLAPMAKALQDNFGVWEGMTTHAYTGDQMIIDGPHRGGLRRAGAGANIVPNSGTAA 238
Qy 328 KAIGLVIPELNGKLDGAAQRPVPTGSVTBELVVTLDKNVSVDEINAAKKAASNDSEGYTE 387
Db 239 KAIGLVIPELNGKLDGSAQRPVPTGSVTBELVAVLEKNVTVDEVNAAKKAASNESYGYTE 298
Qy 388 DPIVSSDIVGVSYGSLFDATQTKMEVDGSQLVKVSWYDNENSYTAQLVRTLEYFAKIA 447
Db 299 DPIVSSDIVGVSYGSLFDATQTKLVLDVKGQVKVSWYDNENSYTAQLVRTLEYFAKIA 358
Qy 448 K 448
Db 359 K 359
```

Search completed: March 29, 2004, 16:45:22
Job time : 100.552 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 16:26:39 ; Search time 26.7463 Seconds
(without alignments)
1611.208 Million cell updates/sec

Title: US-10-650-369-22

Perfect score: 2278

Sequence: 1 MKKITGIILLAVILSAC.....EMSYTAQLVRTLEYFAKIAK 448

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1652.5	72.5	336	2 A42963	glyceralddehyde-3-p
2	1558.5	68.4	336	2 S71350	glyceralddehyde-3-p
3	1535	67.4	335	2 F95235	glyceralddehyde-3-p
4	1535	67.4	359	2 G98099	glyceralddehyde-3-p
5	1357.5	59.6	336	2 F85905	hypothetical prote
6	1280	56.2	337	2 G86694	hypothetical prote
7	1234	54.2	334	2 S34254	glyceralddehyde-3-p
8	1205.5	52.9	336	2 AC1382	glyceralddehyde-3-p
9	1204.5	52.9	336	2 AD1751	glyceralddehyde-3-p
10	1203	52.8	334	2 C96987	glyceralddehyde-3-p
11	1145	50.3	334	2 B82019	glyceralddehyde-3-p
12	1132.5	49.7	336	2 E89850	glyceralddehyde-3-p
13	1132	49.7	334	2 E81001	glyceralddehyde-3-p
14	955.5	41.9	338	2 T96833	glyceralddehyde-3-p
15	941.5	41.3	333	2 F90881	glyceralddehyde-3-p
16	941.5	41.3	333	2 C83737	glyceralddehyde-3-p
17	870.5	38.2	337	2 S73737	glyceralddehyde-3-p
18	861.5	37.8	349	2 F90517	glyceralddehyde-3-p
19	854.5	37.5	337	2 C64233	glyceralddehyde-3-p
20	851	37.4	336	2 A43260	glyceralddehyde-3-p
21	849	37.3	335	1 DES8G	glyceralddehyde-3-p
22	845	37.1	335	1 T2696	glyceralddehyde-3-p
23	836	36.7	335	2 H84094	glyceralddehyde-3-p
24	830	36.4	335	1 DES8GF	glyceralddehyde-3-p
25	820.5	36.0	335	1 A70107	probable glycerald
26	819.5	36.0	333	1 DEHGTT	glyceralddehyde-3-p
27	793.5	34.8	342	2 F70391	glyceralddehyde-3-p
28	790	34.7	336	2 T36020	glyceralddehyde-3-p
29	786.5	34.5	339	2 G70915	glyceralddehyde-3-p

ALIGNMENTS

RESULT 1

A42963

glyceralddehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Streptococcus N/Alternate names: plasmin receptor

C:Species: Streptococcus sp.

C:Date: 10-Mar-1994 #sequence revision 10-Mar-1994 #text_change 03-Jun-2002

C:Accession: A42963; B42963; JH0750

R:Lotenberg, R.; Broder, C.C.; Boyle, M.D.; Kain, S.J.; Schroeder, B.L.; Curtiss III, R

J. Bacteriol. 174, 5204-5210, 1992

A:Title: Cloning, sequence analysis, and expression in Escherichia coli of a streptococ

A:Reference number: A42963; MUID:92355491; PMID:1322883

A:Accession: A42963

A:Molecule type: DNA

A:Residues: 1-336 <LOT>

A:Experimental source: group A, strain 64/14

A:Note: sequence extracted from NCBI backbone (NCBIP:110308)

A:Accession: B42963

A:Molecule type: protein

A:Residues: 2-74,161-164, 'X',166-174,187-211, 'X',213-217 <LO2>

R:Pancholi, V.; Fischetti, V.A.

J. Exp. Med. 176, 415-426, 1992

A:Title: A major surface protein on group A streptococci is a glyceraldehyde-3-phosphate

A:Reference number: JH0750; MUID:92364544; PMID:1500854

A:Accession: JH0750

A:Molecule type: protein

A:Residues: 2-30, 'A',32-40 <PAN>

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: gluconeogenesis; glycolysis; homotetramer; NAD; oxidoreductase

F:152/Active site: Cys #status predicted

Query Match 72.5%; Score 1652.5; DB 2; Length 336;
Best Local Similarity 79.1%; Pred. No. 9.1e-101;
Matches 333; Conservative 2; Mismatches 1; Indels 85; Gaps 1;

Qy	28	MVKVGVNGRIGRIAFRIQNVGVETRIINDLTPNNLAHLKYDVTQGFDTVEV	87
Db	1	MVKVGVNGRIGRIAFRIQNVGVETRIINDLTPNNLAHLKYDVTQGFDTVEV	60
Qy	88	KEGGEFVNGNFIKVSARDPENIDWATGDVEIVLEALEGTVEKGGDFVNGFKIVSAE	147
Db	61	KEGGEFVNGNFIKVSARDPE-----	81
Qy	148	KDPEQIDWATGDVEIVLEIDGTVEKGGFVNGQFKVSAEREPANIDWATGDVEIVLE	207
Db	82	-----NIDWATGDVEIVLE	95
Qy	208	ATGFFAKKAEEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDGTETVIGASCTTN	267
Db	96	ATGFFAKKAEEKHLHANGAKKVVITAPGNDVKTVVFNTHDILDGTETVIGASCTTN	155
Qy	268	CLAPMAKALHDAFGIOKGLMTTHATVGTQMILDGPHRGDLRRARAGAAVFNSTGAA	327

Db 156 CLAPMAKALHDAFGIQKGLMTTTHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 215
 QY 328 KAIGLVIPELNGKLDGAARQVPVPTGVTSLVTLTKVSVDEINAAKAAASNDSPGYTE 387
 Db 216 KAIGLVIPELNGKLDGAARQVPVPTGVTSLVTLTKVSVDEINAAKAAASNDSPGYTE 275
 QY 388 DPVSSDIVGVSGLSFDAQTQKMEVDGSQLVKVSWYDNEMSYTAQLVRLTYFAKIA 447
 Db 276 DPVSSDIVGVSGLSFDAQTQKMEVDGSQLVKVSWYDNEMSYTAQLVRLTYFAKIA 335
 QY 448 K 448
 Db 336 K 336
 RESULT 2
 S71350
 glyceraldhyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Streptococcus
 C:Species: Streptococcus "equisimilis"
 C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 03-Jun-2002
 C:Accession: S71350
 R:Gase, K.; Gase, A.; Schirmer, H.; Malke, H.
 Eur. J. Biochem. 239, 42-51, 1996
 A:Title: Cloning, sequencing and functional overexpression of the Streptococcus equisimilis
 inding protein. Purification and biochemical characterization of the protein.
 A:Reference number: S71350; MUID:96305364; PMID:8706717
 A:Accession: S71350
 A:Molecule type: DNA
 A:Residues: 1-336 <GAS>
 A:Cross-references: EMBL:X97788; NID:g1478268; PIDN:CAA66377.1; PID:g1478269
 A:Experimental source: strain H46A
 C:Genetics:
 A:Gene: gapC
 C:Superfamily: glyceraldhyde-3-phosphate dehydrogenase
 C:Keywords: gluconeogenesis; glycolysis; NAD; oxidoreductase
 F:4-34/Region: beta-alpha-beta NAD nucleotide-binding fold
 F:152/Active site: Cys #status predicted
 Query Match 68.4%; Score 1558.5; DB 2; Length 336;
 Best Local Similarity 76.0%; Pred. No. 1.3e-94;
 Matches 320; Conservative 2; Mismatches 14; Indels 85; Gaps 1;
 QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRPDGTVEV 87
 Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRPDGTVEV 60
 QY 88 KEGFVNGNFIKVSARDPENIDWATDGVETVLEALEGTVKDGDFVNGKFIKVSAAE 147
 Db 61 KEGFVNGNFIKVSARDPE----- 81
 QY 148 KDPEQIDWATDGVETVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVETVLE 207
 Db 82 -----NIDWATDGVETVLE 95
 QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVVNTNHDILDTGTEVTSASCTTN 267
 Db 96 ATGFFAKKEAEKHLHANGAKKVVITAPGNDVKTVVNTNHDILDTGTEVTSASCTTN 154
 QY 268 CLAPMAKALHDAFGIQKGLMTTTHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 327
 Db 156 CLAPMAKALHDAFGIQKGLMTTTHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 214
 QY 328 KAIGLVIPELNGKLDGAARQVPVPTGVTSLVTLTKVSVDEINAAKAAASNDSPGYTE 387
 Db 215 KAIGLVIPELNGKLDGAARQVPVPTGVTSLVTLTKVSVDEINAAKAAASNDSPGYTE 274
 QY 388 DPVSSDIVGVSGLSFDAQTQKMEVDGSQLVKVSWYDNEMSYTAQLVRLTYFAKIA 447
 Db 275 DPVSSDIVGVSGLSFDAQTQKMEVDGSQLVKVSWYDNEMSYTAQLVRLTYFAKIA 334
 QY 448 K 448
 Db 336 K 336
 RESULT 4
 G98099
 glyceraldhyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) [imported] - St
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 03-Jun-2002
 C:Accession: G98099
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Bargett, S.; DeHoff, B.S.; E
 r, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; N
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5703-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

RESULT 3

F95235
 glyceraldhyde 3-phosphate dehydrogenase [imported] - Streptococcus pneumoniae (strain 1
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heic
 on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
 nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: F95235
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-335 <KUR>
 A:Cross-references: GB:AB005672; PIDN:AAK76079.1; PID:g14973522; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP2012
 C:Superfamily: glyceraldhyde-3-phosphate dehydrogenase

Query Match 67.4%; Score 1535; DB 2; Length 335;
 Best Local Similarity 73.6%; Pred. No. 4.4e-93;
 Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRPDGTVEV 87
 Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETVTRINDLTPNMLAHLKYDTTQGRPDGTVEV 60
 QY 88 KEGFVNGNFIKVSARDPENIDWATDGVETVLEALEGTVKDGDFVNGKFIKVSAAE 147
 Db 61 KEGFVNGNFIKVSARDPE----- 81
 QY 148 KDPEQIDWATDGVETVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVETVLE 207
 Db 82 -----QIDWATDGVETVLE 95
 QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVVNTNHDILDTGTEVTSASCTTN 267
 Db 96 ATGFFAKKEAEKHLHANGAKKVVITAPGNDVKTVVNTNHDILDTGTEVTSASCTTN 154
 QY 268 CLAPMAKALHDAFGIQKGLMTTTHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 327
 Db 155 CLAPMAKALQDNFVGVVGLMTTTHAYTGDMILDGPHRGDLRRARAGAAIVPNSTGAA 214
 QY 328 KAIGLVIPELNGKLDGAARQVPVPTGVTSLVTLTKVSVDEINAAKAAASNDSPGYTE 387
 Db 215 KAIGLVIPELNGKLDGAARQVPVPTGVTSLVTLTKVSVDEINAAKAAASNDSPGYTE 274
 QY 388 DPVSSDIVGVSGLSFDAQTQKMEVDGSQLVKVSWYDNEMSYTAQLVRLTYFAKIA 447
 Db 275 DPVSSDIVGVSGLSFDAQTQKMEVDGSQLVKVSWYDNEMSYTAQLVRLTYFAKIA 334
 QY 448 K 448
 Db 335 K 335

RESULT 4

G98099
 glyceraldhyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) [imported] - St
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 03-Jun-2002
 C:Accession: G98099
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Bargett, S.; DeHoff, B.S.; E
 r, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; N
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5703-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: G98099
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-359 <KUR>
A:Cross-references: GB:AE007317; PIDN:AA00628.1; PID:g15459513; GSPDB:GN00174
C:Genetics:
A:Gene: gapA
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: oxidoreductase

Query Match 67.4%; Score 1535; DB 2; Length 359;
Best Local Similarity 73.6%; Pred. No. 4.8e-93;
Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;

QY 28 MVKVGINGFRIGRLAFRRIQNVGEVETRLNDLTPNMLAHLKYDTTQGRFDGTV 87
DB 25 MVKVGINGFRIGRLAFRRIQNVGEVETRLNDLTPNMLAHLKYDTTQGRFDGTV 84
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVKDGDFVNGKFIKVS 147
DB 85 KEGGFVNGKFIKVSARDPE-----QIDWATDGVIEVLE 105
QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAREPANIDWATDGVIEV 207
DB 106 ----- 207
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTIVFNTNHDILDTGTVISGASCTTN 267
DB 120 ATGFFAKKEAAEKHL-KGGAKKVVITAPGNDVKTIVFNTNHDVLDGTETVISGASCTTN 178
QY 268 CLAPMAKALHDAPFGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
DB 179 CLAPMAKALQDNFVGVVGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 238
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVVTLTKNVSVDENAAKKAASNDSPGYTE 387
DB 239 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVVTLTKNVSVDENAAKKAASNDSPGYTE 298
QY 388 DPVSSDIVGVSGLFDATQTKMVEVDGSQLVKVSWYDNMSYTAQLVRLTLEYFAKIA 447
DB 299 DPVSSDIVGVSGLFDATQTKVLDVGKLVKVSWSYDNMSYTAQLVRLTLEYFAKIA 358
QY 448 K 448
DB 359 K 359

RESULT 5
F86905
hypothetical protein gapB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: F86905
R:Boletoin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: F86905
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-336 <STO>
A:Cross-references: GB:AE005176; PID:g12723446; PIDN:AAK06344.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: gapB
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 59.6%; Score 1357.5; DB 2; Length 336;
Best Local Similarity 65.1%; Pred. No. 1.8e-81;
Matches 274; Conservative 21; Mismatches 41; Indels 85; Gaps 1;

QY 28 MVKVGINGFRIGRLAFRRIQNVGEVETRLNDLTPNMLAHLKYDTTQGRFDGTV 87
DB 1 MVKVGINGFRIGRLAFRRIQNVGEVETRLNDLTPNMLAHLKYDTTQGRFDGTV 54
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVKDGDFVNGKFIKVS 147
DB 55 ----- 54
QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAREPANIDWATDGVIEV 207
DB 55 -----DKVEVKGDFVNGKFIKVSAREPANINAEVAGAEIVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTIVFNTNHDILDTGTVISGASCTTN 267
DB 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGSDVKTIVFNTNHDVLDGTETVISGASCTTN 155
QY 268 CLAPMAKALHDAPFGIQKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
DB 156 CLAPMAKALQDNFVGVVGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVVTLTKNVSVDENAAKKAASNDSPGYTE 387
DB 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVVTLTKNVSVDENAAKKAASNDSPGYTE 275
QY 388 DPVSSDIVGVSGLFDATQTKMVEVDGSQLVKVSWYDNMSYTAQLVRLTLEYFAKIA 447
DB 276 DPVSSDIVGVSGLFDATQTEVTSADGALVKVSWYDNMSYTNLVRLLAYFAKIA 335
QY 448 K 448
DB 336 K 336

RESULT 6
G86694
hypothetical protein gapA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: G86694
R:Boletoin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: G86694
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <STO>
A:Cross-references: GB:AE005176; PID:g12723446; PIDN:AAK04657.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: gapA
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 56.2%; Score 1280; DB 2; Length 337;
Best Local Similarity 61.4%; Pred. No. 2e-76;
Matches 259; Conservative 31; Mismatches 46; Indels 86; Gaps 2;

QY 28 MVKVGINGFRIGRLAFRRIQNVGEVETRLNDLTPNMLAHLKYDTTQGRFDGTV 87
DB 1 MVKVGINGFRIGRLAFRRIQNVGEVETRLNDLTPNMLAHLKYDTTQGRFDGTV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVKDGDFVNGKFIKVS 147
DB 61 KEGGFVNGKFIKVSARDPE-----DIQWADSGVIEVLE 81
QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAREPANIDWATDGVIEV 207
DB 82 ----- 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTIVFNTNHDILDTGTVISGASCTTN 267
DB 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTIVFNTNHDVLDGTETVISGASCTTN 155

QY 268 CLAPMAKALHDARGIOKGLMTTHAYTGDMILDGPHRGGLRPARAGANIVPNSGAA 327
 DB 156 SLAPMADALNKNFVGKGTWTHSYTGDMTLDGPHRGGLRPARAGANIVPNSGAA 215
 QY 328 KAIGLVIPPELNGKLDGAAQRPVPTGVTBLVTLTKNVSVDEINAAKAAASNSDFGYTE 387
 DB 216 KAIGLVIPPELNGKLDGAAQRPVPTGVTBLVTLTKNVSVDEINAAKAAASNSDFGYTE 275
 QY 388 DPTVSSDIVGVSGLSFDAQTQKMEV-DGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKI 446
 DB 276 DEIVSSDIIMGAYGSLFDATLITVTLKDGGLVKTAAWYDNEMSYTAQLVRLTLEYFAKI 335
 QY 447 AK 448
 DB 336 AK 337

RESULT 7

S34254
 glyceralddehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Clostridium F
 C:Species: Clostridium pasteurianum
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 03-Jun-2002
 C:Accession: S34254
 R:Oster, I.; Asoohi, O.; Scherrer, S.; Branlant, G.; Branlant, C.
 submitted to the EMBL Data Library, May 1993
 A:Description: Nucleotide sequence of the glyceralddehyde-3-phosphate dehydrogenase from
 A:Reference number: S34254
 A:Accession: S34254
 A:Molecule type: DNA
 A:Residues: 1-334 <OST>
 A:Cross-references: EMBL:X72219; NID:g311923; PIDN:CAA51020.1; PID:g311924
 C:Superfamily: Glyceralddehyde-3-phosphate dehydrogenase
 C:Keywords: oxidoreductase

Query Match 54.2%; Score 1234; DB 2; Length 334;
 Best Local Similarity 59.3%; Pred. No. 2.1e-73;
 Matches 249; Conservative 34; Mismatches 51; Indels 86; Gaps 3;
 QY 29 VVKVINGFGFRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVGVK 88
 DB 1 MTKVAINGFGFRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVGVK 60
 QY 89 EGGFEVNGNFTKVSABRDPENIDWATDGVLEALEGTVGVKDGDFVNGKFIKVSABK 148
 DB 61 EGAFVNGKVKVFAEADPEKLPV-----GEL-----GIDVVLEK 95
 QY 149 DPEQIDWATDGVLEIDGTVGVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEA 208
 DB 85 -----GEL-----GIDVVLEK 95

QY 209 TSFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVPNTNHDILDTGTVTISGASCTTNC 268
 DB 96 TGFTTKKEAAHVRRA-GAKKVVISAPAGNDLKTIVFNVNNDLDTGTVTISGASCTTNC 154
 QY 269 LAPMAKALHDARGIOKGLMTTHAYTGDMTLDGPHRGGLRPARAGANIVPNSGAAK 328
 DB 155 LAPMAKALHDARGIOKGLMTTHAYTGDMTLDGPHRGGLRPARAGANIVPNSGAAK 214
 QY 329 AIGLVIPPELNGKLDGAAQRPVPTGVTBLVTLTKNVSVDEINAAKAAASNSDFGYTE 388
 DB 215 AIAQVIPLELNGKLDGAAQRPVPTGVTBLVTLTKNVSVDEINAAKAAASNSDFGYTE 274
 QY 389 PTVSSDIVGVSGLSFDAQTQKMEV-DGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKI 448
 DB 275 EIVSADVGVISGLSFDAQTQKMEV-DGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKI 334

RESULT 8

AC1382
 glyceralddehyde 3-phosphate dehydrogenase homolog gap [imported] - Listeria monocytogenes
 C:Species: Listeria monocytogenes
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 C:Accession: AC1382

R:Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Me
 ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
 A:Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AC1382
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-336 <GLA>
 A:Cross-references: GB:NC.003210; PIDN:CAD00537.1; PID:g16411947; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: gap
 C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase

Query Match 52.9%; Score 1205.5; DB 2; Length 336;
 Best Local Similarity 58.5%; Pred. No. 1.5e-71;
 Matches 247; Conservative 34; Mismatches 54; Indels 87; Gaps 4;
 QY 28 MVVKVINGFGFRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVGVK 87
 DB 1 MTKVINGFGFRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDGTVGVK 60
 QY 88 KEGFEVNGNFTKVSABRDPENIDWATDGVLEALEGTVGVKDGDFVNGKFIKVSABK 147
 DB 61 HDGFEVNGKVKVLANENPSELPMG-----DL----- 88
 QY 148 KPEQIDWATDGVLEIDGTVGVKEGGFEVNGQFVKVSAEREPANIDWATDGVLEA 207
 DB 89 -----GVDIVLE 95

QY 208 ATFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVPNTNHDILDTGTVTISGASCTTNC 267
 DB 96 CTGFTTAKDAELHIKA-GAKKVVISAPAGNDLKTIVFNVNNDLDTGTVTISGASCTTNC 154
 QY 268 CLAPMAKALHDARGIOKGLMTTHAYTGDMTLDGPHRGGLRPARAGANIVPNSGAA 327
 DB 155 CLAPMAKALHDARGIOKGLMTTHAYTGDMTLDGPHRGGLRPARAGANIVPNSGAA 214
 QY 328 KAIGLVIPPELNGKLDGAAQRPVPTGVTBLVTLTKNVSVDEINAAKAAASNSDFGYT 386
 DB 215 KAIGLVIPLELNGKLDGAAQRPVPTGVTBLVTLTKNVSVDEINAAKAAASNSDFGYT 274
 QY 387 EPTVSSDIVGVSGLSFDAQTQKMEV-DGSQLVKVSVWYDNEMSYTAQLVRLTLEYFAKI 446
 DB 275 SQVWSSDIIMGAYGSLFDATLITVTLKDGGLVKTAAWYDNEMSYTAQLVRLTLEYFAKI 334

QY 447 AK 448
 DB 335 AK 336

RESULT 9

AD1751
 glyceralddehyde 3-phosphate dehydrogenase homolog gap [imported] - *Listeria innocua* (str
 C:Species: *Listeria innocua*
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 C:Accession: AD1751
 R:Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Me
 ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
 A:Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AD1751
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-336 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97780.1; PID:g16415075; GSPDB:GN00178
A:Experimental source: strain Clip11262

C:Genetics:

A:Gene: gap

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 52.9%; Score 1204.5; DB 2; Length 336;
Best Local Similarity 59.5%; Pred. No. 1.8e-71;
Matches 247; Conservative 33; Mismatches 55; Indels 87; Gaps 4;

QY 28 MVKVGNGFGRIGRLAFRIQNVGVETRIINDLTPNMLHLKYDTTQGRFDGTV 87

Db 1 MTVKVGNGFGRIGRLAFRIQNVGVETRIINDLTPNMLHLKYDTTQGRFDGTV 60

QY 88 KEGFVNGNFIKVSABERDPENIDWATDGVIEVLEAGTGVKDGDFVNGKFIKVS 147

Db 61 HDGFVNGNFKVLANRNPPELPWG-----DL----- 88

QY 148 KDFEQIDWATDGVIEVLEIDGTVEVEKGGFVNGQFVKVSAERPNIDWATDGVIEV 207

Db 89 -----GVDIVLE 95

QY 208 ATGFFAKKAAEKHLHANGAKKVVITAPGNDVKTIVFNTNHDLDGTETVIGASCTTN 267

Db 96 CTGFFTAQDKAELHXA-GAKKVISAPATGDKTIVVNNHETLDGTETVIGASCTTN 154

QY 268 CLAPMAKALHDAGIQKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAAIVPNSTGAA 327

Db 155 CLAPMAKVLDEKFGVVEGLMTTHAYTGDQNTLDAPHPKDFRRARAAENIIPNTGAA 214

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSLTELVTLDKGVSVDEINAAKAAAS-DSFGYT 386

Db 215 KAIGLVIPTELKGLDGAORVPVPTGSLTELVTLDKGVSVDEINAAKAAASDPETFGYT 274

QY 387 EDIVSSDIVGVSGSLFDATQTKWMEVDGSQLVKKVSVYDNMSYTAQLVRLTLEYFAKI 446

Db 275 SDQWSSDIKGMTFGSLFDEDTQTKVLTVDGQQLVKTAVMYDNMSYTAQLVRLTLEYFAKI 334

QY 447 AK 448

Db 335 AK 336

RESULT 10

Cy96987

glyceraldehyde 3-phosphate dehydrogenase, gene gapC [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001

C:Accession: C96987

R:Rolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: C96987

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-334 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK78686.1; PID:g15023589; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC0709

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 52.8%; Score 1203; DB 2; Length 334;

Best Local Similarity 57.4%; Pred. No. 2.2e-71;

Matches 241; Conservative 41; Mismatches 52; Indels 86; Gaps 3;

QY 29 VVKVGNGFGRIGRLAFRIQNVGVETRIINDLTPNMLHLKYDTTQGRFDGTV 88

Db 1 MAKIANGFGRIGRLALRIEVLPGLEVAINDLTDKMLHLFKYDSSQGRNGEIEVK 60

QY 89 EGGFVNGNFIKVSABERDPENIDWATDGVIEVLEAGTGVKDGDFVNGKFIKVS 148

Db 61 EGAFVNGNFKVFAEADPEKLPWG-----DL----- 87

QY 149 DPEQIDWATDGVIEVLEIDGTVEVEKGGFVNGQFVKVSAERPNIDWATDGVIEV 208

Db 88 -----GIDVILEC 95

QY 209 TSFFAKKAAEKHLHANGAKKVVITAPGNDVKTIVFNTNHDLDGTETVIGASCTTN 268

Db 96 TGFFTKKAAEAHVA-GAKKVISAPAGNDLKIIVFVNNHEDLDGTETVIGASCTTN 154

QY 269 LAPMAKALHDAGIQKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAAIVPNSTGAA 328

Db 155 LAPMAKVLNDFKIEKFGMTTHAYTGDQNTLDGPHRGDLRRARAAVSIIPNSTGAA 214

QY 329 AIGLVIPELNGKLDGAAQRPVPTGSLTELVTLDKGVSVDEINAAKAAASDSFGYT 388

Db 215 ALSQVIPDLAGLDGNAQRPVPTGSLTELVTLDKGVSVDEINAAKAAADESFGYT 274

QY 389 PIVSSDIVGVSGSLFDATQTKWMEVDGSQLVKKVSVYDNMSYTAQLVRLTLEYFAKI 448

Db 275 PIVSADVGVNGSLFDATLTKIVDNGSQLVKTAAWYDNMSYTSQVRLTLEYFAKI 334

RESULT 11

B82019

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) C NMA0246 [impc

C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002

C:Accession: B82019

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

R:Hoyle, S.; Jags, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A:Reference number: A81775; MUID:2022556; PMID:10761919

A:Accession: B82019

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-334 <PAR>

A:Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB83554.1; PID:g737900

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA0246

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: oxidoreductase

Query Match 50.3%; Score 1145; DB 2; Length 334;

Best Local Similarity 56.4%; Pred. No. 1.4e-67;

Matches 235; Conservative 35; Mismatches 61; Indels 86; Gaps 4;

QY 28 MVKVGNGFGRIGRLAFRIQNVGVETRIINDLTPNMLHLKYDTTQGRFDGTV 87

Db 1 MSKVAINGFGRIGRLALRIEVLPGLEVAVNDLTPAEMLLHLFKYDSTQGRF----- 54

QY 88 KEGFVNGNFIKVSABERDPENIDWATDGVIEVLEAGTGVKDGDFVNGKFIKVS 147

Db 55 -----QGTALDKDAIVNGKEIKVFAN 77

QY 148 KDFEQIDWATDGVIEVLEIDGTVEVEKGGFVNGQFVKVSAERPNIDWATDGVIEV 207

Db 78 PNPEELPW-----GEL-----GVDVILE 95

QY 208 TSFFAKKAAEKHLHANGAKKVVITAPGNDVKTIVFNTNHDLDGTETVIGASCTTN 267

Db 96 CTGFFTKKAAEHXA-GAKKVISAPAGNDLKIIVFVNNHEDLDGTETVIGASCTTN 154

QY 268 CLAPMAKALHDAGIQKGLMTTHAYTGDQMIIDGPHRGDLRRARAGAAIVPNSTGAA 327

Db 155 CLAPMAVLOKEFGVVEGLMTTHAYTGDQNTLDAPHPKDFRRARAAALNIVPNSTGAA 214

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGSLTELVTLDKGVSVDEINAAKAAASDSFGYT 387

Db 215 KAIGLVIPELNGKLDGAAQRPVPTGSLTELVTLDKGVSVDEINAAKAAASESYNE 274

QY 388 DPIVSSDIVGVSGLSDATQTKMEYDGSQLVKVSVDNEMSYTAQLVRLTYEFA 444
Db 275 DQIVSSDVVGIEGSLFDATQTRVMTVGKQLVKTAVWYDNEMSYTCQLVRLTYEFA 331

RESULT 12
E89850
glyceraldehyde-3-phosphate dehydrogenase [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: E89850
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogino, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: E89850
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-336 <KUR>
A:Cross-references: GB:BA000018; PID:g13700663; PIDN:BA041960.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: gap
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 49.7%; Score 1132.5; DB 2; Length 336;
Best Local Similarity 54.7%; Pred. No. 8.9e-67;
Matches 231; Conservative 44; Mismatches 60; Indels 87; Gaps 4;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDTPNMLAHLKYDTTQGRPDGTVEV 87
Db 1 NAVKVAINGFGRIGRLAFRRIOEVEGLEWAVNDLTDMLAHLKYDTTQGRFTGEV 60

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIELEALGTVKDGDFVNGKFIKVS 147
Db 61 VDGFVRNGKEVKSFPDASKLPK-----DUN----- 89

QY 148 KPQEQIDWATDGVIEVLEIDGTVEKGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 90 -----IDVLE 95

QY 208 ATSPFAKAEAEKHLHANGAKKVVITAPGNDVKTVPNTNHDILDTGTEVTSASCTTN 267
Db 96 CTGFYTDKKAQAHEA-GAKKVLISAPATGDLKTIIVNTNHOELDGSFVTSASCTTN 154

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQIMLDGPHRGDLRRARAGAAIVPNSTGAA 327
Db 155 SLAPVAKVNDPGLVEGLMTTHAYTGDQNTQDAPHRKGGKRRARAAENIIPNSTGAA 214

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLDK-NVSVDEINAKMAANDSFGYT 386
Db 215 KAIGKVIPEIDGKLDGGAQRVPVATGSLTSLTVVLEKQDVTVEQVNEAMKASNESFGYT 274

QY 387 EDPVSSDIVGVSGLSDATQTKMEYDGSQLVKVSVDNEMSYTAQLVRLTYEFAKI 446
Db 275 EDEIVSSDVVGNTGSLFDATQTRVMSVGRQLVKTAVWYDNEMSYTAQLVRLTYEFAEL 334

QY 447 AK 448
Db 335 SK 336

RESULT 13
E81001
glyceraldehyde 3-phosphate dehydrogenase NMB2159 [imported] - Neisseria meningitidis (strain NMB2159)
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: E81001
R:Testelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V.
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; PMID:20175755; PMID:10710307
A:Accession: E81001
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <TET>
A:Cross-references: GB:AE002563; GB:AE002098; NID:G7227405; PIDN:AAF42467.1; PID:G722741
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB2159
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

Query Match 49.7%; Score 1132; DB 2; Length 334;
Best Local Similarity 55.6%; Pred. No. 9.6e-67;
Matches 232; Conservative 37; Mismatches 66; Indels 86; Gaps 4;

QY 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDTPNMLAHLKYDTTQGRPDGTVEV 87
Db 1 MSIKVAINGFGRIGRLALRQIEKAHDIEVAVNDLTPAEMLLHLFKYDSTQGRF----- 54

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALGTVKDGDFVNGKFIKVS 147
Db 55 -----QGTAEKDDDAIVNGKEIKVFAN 77

QY 148 KPQEQIDWATDGVIEVLEIDGTVEKGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 78 PNPEELPW-----GEL-----GVDVILE 95

QY 208 ATSPFAKAEAEKHLHANGAKKVVITAPGNDVKTVPNTNHDILDTGTEVTSASCTTN 267
Db 96 CTGFYTDKKAQAHEA-GAKKVLISAPGNDVKTVPNTNHDILDTGTEVTSASCTTN 154

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQIMLDGPHRGDLRRARAGAAIVPNSTGAA 327
Db 155 CLAPMAVILQREFGVVEGLMTTHAYTGDQNTDAPHRKGGDLRRARAAALIVPNSTGAA 214

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTSLVTLVTLDKKNVSVDEINAKMAANDSFGYTE 387
Db 215 KAIGLVIPELNGKLDGGAQRVPVATGSLTSLTVLRLERVTKEEINAKMAAASISYGN 274

QY 388 DPIVSSDIVGVSGLSDATQTKMEYDGSQLVKVSVDNEMSYTAQLVRLTYEFA 444
Db 275 DQIVSSDVVGIEGSLFDATQTRVMTVGKQLVKTAVWYDNEMSYTCQLVRLTYEFA 331

RESULT 14
T09633
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Lactobacillus
C:Species: Lactobacillus delbrueckii
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T09633
R:Branny, P.; Delatorre, F.; Garel, J.R.
Microbiology 144, 905-914, 1998
A:Title: An operon coding for three glycolytic enzymes in *Lactobacillus delbrueckii* subsp.
A:Reference number: Z16788; PMID:98240227; PMID:9579064
A:Accession: T09633
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-338 <BRA>
A:Cross-references: EMBL:AJ000339; NID:G2624189; PIDN:CAA04014.1; PID:G2624191
A:Experimental source: subsp. bulgaricus, strain B107
C:Genetics:
A:Gene: gap
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: gluconogenesis; glycolysis; oxidoreductase

Query Match 41.9%; Score 955.5; DB 2; Length 338;
Best Local Similarity 47.3%; Pred. No. 3.3e-55;
Matches 200; Conservative 46; Mismatches 88; Indels 89; Gaps 4;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 15:59:33 ; Search time 19.1045 Seconds

(without alignments)
1221.045 Million cell updates/sec

Title: US-10-650-369-22

Perfect score: 2278

Sequence: 1 MKKITGIILLALLVILSAC.....EMSYTAQLVRLTLEYPKIAK 448

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	1650.5	72.5	335	1	G3P_STRPY P50467 streptococc
2	1646.5	72.3	335	1	G3P_STRP3 Q8k8m9 streptococc
3	1553.5	68.2	335	1	G3P_STRPQ Q59906 streptococc
4	1280	56.2	337	1	G3P_LACLA P52987 lactococcus
5	1234	54.2	334	1	G3P_CLOPA O59309 clostridium
6	1203	52.8	334	1	G3P_CLOAB O52631 clostridium
7	1138.5	50.0	336	1	G3P_STAEP Q8cpys staphylococ
8	1132.5	49.7	336	1	G3P_LACDE Q925c5 staphylococ
9	955.5	41.9	338	1	G3P_STAAM O32755 lactobacill
10	941.5	41.3	333	1	G3P_BCO57 P58072 escherichia
11	940.5	41.3	333	1	G3P_ECOLI P33898 escherichia
12	870.5	38.2	337	1	G3P_MYCPN P75358 mycoplasma
13	854.5	37.5	337	1	G3P_MYCGE P47543 mycoplasma
14	850	37.3	334	1	G3P_CORGL Q01651 corynebacte
15	844	37.1	334	1	G3P_BACSU P09124 bacillus su
16	840	36.9	334	1	G3P_BACME P23722 bacillus me
17	825	36.2	334	1	G3P_BACST P00362 bacillus st
18	820.5	36.0	335	1	G3P_BORBU P46795 borrelia bu
19	818.5	35.9	332	1	G3P_THEME P17721 thermotoga
20	793.5	34.8	342	1	G3P_AQUAE O67161 aquifex aeo
21	790	34.7	336	1	G3P_STRCO Q92518 streptomyce
22	786.5	34.5	339	1	G3P_MYCTU O06822 mycobacteri
23	783.5	34.4	339	1	G3P_MYCLE P46713 mycobacteri
24	777.5	34.1	339	1	G3P_MYCAV P94915 mycobacteri
25	763.5	33.5	330	1	G3P_SALTY P24165 salmonella
26	759.5	33.3	333	1	G3P_SALTY P06977 escherichia
27	754	33.1	339	1	G3P_REIN P44304 haemophilus
28	743.5	32.6	330	1	G3P_TRYBB P10097 trypanosoma
29	740	32.5	332	1	G3P_RALSO P52694 raietonia s
30	739	32.4	332	1	G3P_STRAU Q59800 streptomyce
31	737	32.4	337	1	G3P_RHRA Q9c136 rhizomucor
32	737	32.4	337	1	G3P_MONAN P53430 monascus an
33	734	32.2	336	1	G3P_SCHPO P78958 schizosacch

ALIGNMENTS

RESULT 1

ID	G3P_STRPY	STANDARD;	PRT;	335 AA.
AC	P50467;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)			
DE	(Plasminogen-binding protein) (Plasmin receptor).			
GN	GAP OR PLR OR GAPA OR SPY0274 OR SPYM18_0261.			
OS	Streptococcus pyogenes, and			
OS	Streptococcus pyogenes (serotype M18).			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus			
OX	NCBI_TaxID=1314, 186103;			
RN	SEQUENCE FROM N.A., AND SEQUENCE OF 1-73; 160-173 AND 186-216.			
RP	STRAIN=64/14;			
RX	MEDLINE=92355491; PubMed=1322883;			
RA	Lottenberg R., Broder C.C., Boyle M.D., Kain S.J., Schroeder B.L.,			
RA	Curtiss R. III;			
RT	"Cloning, sequence analysis, and expression in Escherichia coli of a			
RT	streptococcal plasmin receptor.";			
RL	J. Bacteriol. 174:5204-5210(1992).			
[2]				
RN	SEQUENCE FROM N.A.			
RP	STRAIN=SF370 / ATCC 700294 / Serotype M1;			
RX	MEDLINE=21192684; PubMed=11296296;			
RA	Ferrecci J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,			
RA	Primeaux C., Szate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,			
RA	Yuan Y., Jia H.G., Najaf F.Z., Ren Q., Zhu H., Song L., White J.,			
RA	Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;			
RT	"Complete genome sequence of an M1 strain of Streptococcus			
RT	pyogenes";			
EL	Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).			
[3]				
RN	SEQUENCE FROM N.A.			
RP	STRAIN=MGAS8232 / Serotype M18;			
RX	MEDLINE=21927593; PubMed=11917108;			
RA	Smoot J.C., Barbian K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,			
RA	Sylva G.L., Sturdevant D.E., Ricklefs S.W., Porcella S.F., Zhang Q.,			
RA	Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,			
RA	Kapur V., Daly J.A., Veasy L.G., Musser J.M.;			
RT	"Genome sequence and comparative microarray analysis of serotype M18			
RT	group A Streptococcus strains associated with acute rheumatic fever			
RT	outbreaks.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).			
[4]				
RN	SEQUENCE OF 1-11; 20-30; 103-128; 162-171 AND 199-215.			
RP	STRAIN=URS4 / Serotype M6;			
RA	Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,			
RA	VanBogelen R.A.;			
RT	"Two-dimensional gel electrophoresis map of Streptococcus pyogenes			
RT	proteins.";			
RL	Submitted (MAY-2000) to Swiss-Prot.			
CC	-!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.			

34	732.5	32.2	337	1	G3P2_ANASP	P58554 anabaena sp
35	729	32.0	337	1	G3P_CLAPU	Q00594 claviceps p
36	725	31.8	331	1	G3P_THERAQ	P00361 thermus aqu
37	724.5	31.8	336	1	G3P2_SYNY3	P80505 synecocyst
38	723	31.7	338	1	G3P_NEUCR	P54118 neurospora
39	722	31.7	337	1	G3P_COCURJ	P29497 cochlidiobol
40	721	31.7	337	1	G3P_COCURJ	P28844 curvularia
41	719.5	31.6	333	1	G3P_STRAE	P54226 streptomyce
42	717.5	31.5	330	1	G3P2_LEIME	O01558 leishmania
43	717	31.5	335	1	G3P2_SCHPO	O43026 schizosacch
44	715	31.4	336	1	G3P_ASPNG	Q12552 aspergillus
45	713	31.3	336	1	G3P_EMENT	P20445 emericella

CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SPIN: 1.2.1.12 (GAPDH)
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M95569; AAA26953.1; -
CC EMBL; AE006494; AAK33348.1; -
CC EMBL; AE009973; AAL97041.1; -
CC HSP; P00362; IGD1.
CC InterPro; IPR000173; GAP_dhhdhgenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC PRINTS; PR00078; G3PDHGRGNASE.
CC PROSITE; PS00071; GAPDH; 1.
CC GlycoLysis; Oxidoreductase; NAD; Complete proteome.
CC INIT_MET 0
CC BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY
CC ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
CC CONFLICT 261 261 A -> S (IN REF. 1).
CC SEQUENCE 335 AA; 35811 MW; F06006E253C8A3F CRC64;
Query Match 72.5%; Score 1650.5; DB 1; Length 335;
Best Local Similarity 79.3%; Pred. No. 1.2e-93;
Matches 333; Conservative 1; Mismatches 1; Indels 85; Gaps 1;
QY 29 VVKVINGFGRIGRIFRRIQNVGVETRIINDLTDNMLAHLKYDTTQGRFDGTVGVK 88
Db 1 VVKVINGFGRIGRIFRRIQNVGVETRIINDLTDNMLAHLKYDTTQGRFDGTVGVK 60
QY 89 EGGFEVNGNFIKVAERDPENIDWATDGEIVLEALEGTVEVKGDPVNGKFIKVAEK 148
Db 61 EGGFEVNGNFIKVAERDPE----- 80
QY 149 DPEQIDWATDGEIVLEIDGTVEVKGEGFVNGQFKVSAEREPANIDWATDGEIVLEA 208
Db 81 -----NIDWATDGEIVLEA 95
QY 209 TSFFAKKEAEKHLHANGAKKVVITAPGNDVKVTWNTNHDLDGTETVIGASCCTTNC 268
Db 96 TGFFAKKEAEKHLHANGAKKVVITAPGNDVKVTWNTNHDLDGTETVIGASCCTTNC 155
QY 269 LAPAKALHDFAFGTQGLMTTHIATYGTQDMLDGFHGGDLRARAAGANIVNSTGAAK 328
Db 156 LAPAKALHDFAFGTQGLMTTHIATYGTQDMLDGFHGGDLRARAAGANIVNSTGAAK 215
QY 329 AIGLVIPELNGKLDGAAQRPVPTGVTVELVTLVDKNVSVDEINAAKASNDSPGYTED 388
Db 216 AIGLVIPELNGKLDGAAQRPVPTGVTVELVTLVDKNVSVDEINAAKASNDSPGYTED 275
QY 389 PIVSSDIVGVSGSLFATQTKNVEVDGSQLKVVSVYDNEMSYTAQVLTLYFPKIAK 448
Db 276 PIVSSDIVGVSGSLFATQTKNVEVDGSQLKVVSVYDNEMSYTAQVLTLYFPKIAK 335
RESULT 2
G3P_STRP3
ID - G3P_STRP3
AC Q8K6M9; STANDARD; PRT; 335 AA.

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)
DE (Plasminogen-binding protein) (Plasmin receptor).
GN GAP OR PLR OR SPYM3_0201 OR SP50207.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=198466;
RN [1]_SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133308; PubMed=12122206;
RA Beres S.B., Sylva G.B., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phase-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [2]_SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RX MEDLINE=22683278; PubMed=13799345;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution.";
RL Genome Res. 13:1042-1055(2003).
CC -!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN (By similarity).
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE014140; AAM78808.1; -
CC EMBL; AP005141; BAC63302.1; -
CC InterPro; IPR000173; GAP_dhhdhgenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC PRINTS; PR00078; G3PDHGRGNASE.
CC TIGRfam; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
CC GlycoLysis; Oxidoreductase; NAD; Complete proteome.
CC INIT_MET 0
CC BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE (BY
CC ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS
CC SEQUENCE 335 AA; 35841 MW; 4DCB76F382F6F99 CRC64;
Query Match 72.3%; Score 1646.5; DB 1; Length 335;
Best Local Similarity 79.0%; Pred. No. 2.1e-93;
Matches 332; Conservative 1; Mismatches 2; Indels 85; Gaps 1;
QY 29 VVKVINGFGRIGRIFRRIQNVGVETRIINDLTDNMLAHLKYDTTQGRFDGTVGVK 88


```

Db 1 VVKVINGFGRIGRLAFRRIONIEGVETRIINDLTDPNMLAHLKYDITQGRFDGTVK 60
QY 89 EGGFEVNGNFIKVSAREDPENIDWATDGVIEVLEALEGTVKDGDFVNGKFIKVSAREK 148
Db 61 EGGFEVNGNFIKVSAREDPENIDWATDGVIEVLEALEGTVKDGDFVNGKFIKVSAREK 80
QY 149 DPEQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLEA 208
Db 81 -----NIDWATDGVIEVLEA 95
QY 209 TSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVPNTNHDILDTGTVISGASCTTNC 268
Db 96 TGFPAKKEAEKHLHANGAKKVVITAPGNDVKTVPNTNHDILDTGTVISGASCTTNC 155
QY 269 LAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAANIVPNSGAAK 328
Db 156 LAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAANIVPNSGAAK 215
QY 329 AIGLVIPELNGKLDGAAQRPVPTGTVTLVTLTKVNSVDEINAAKAAASNDSEGYTED 388
Db 216 AIGLVIPELNGKLDGAAQRPVPTGTVTLVTLTKVNSVDEINAAKAAASNDSEGYTED 275
QY 389 PIVSSDIVGVSGLFDATQTKMVEDGSQLVKVSWYDNEMSYTAOLVRLTLEYFAKIAK 448
Db 276 PIVSSDIVGVSGLFDATQTKMVEDGSQLVKVSWYDNEMSYTAOLVRLTLEYFAKIAK 335

RESULT 3
G3P_STREQ STANDARD; PRT; 335 AA.
AC Q5906;
AT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH)
DE (Plasminogen-binding protein) (Plasmin receptor).
GN GAP OR GAPC.
OS Streptococcus equisimilis
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H46A;
RX MEDLINE=96305364; PubMed=8706717;
RA Gase K., Gase A., Schirmer H., Malke H.;
RT "Cloning, sequencing and functional overexpression of the
RT Streptococcus equisimilis H46A gapC gene encoding a
RT glyceraldehyde-3-phosphate dehydrogenase that also functions as a
RT plasminogen-binding protein. Purification and biochemical
RT characterization of the protein.";
RL Eur. J. Biochem. 239:42-51(1996).
CC -!- FUNCTION: ALSO BINDS HUMAN PLASMINOGEN.
CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -!- PATHWAY: Second phase of glycolysis; first step.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X97788; CAA66377.1; -
CC EMBL; Y12602; CAA73174.1; -
CC F01; S71350; S71350.
CC HSSP; P00362; IGD1.

```

```

DR InterPro; IPR000173; GAP dhhydrogenase.
DR InterPro; IPR006424; GAPDH-1.
DR Pfam; PF00044; Spdh_1.
DR Pfam; PF02800; Spdh_C; 1.
DR PRINTS; PRO0078; G3PDHGRNASE.
DR TIGRfams; TIGR01534; GAPDH-1; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD.
FT INIT MET 0
FT BINDING 151 151
FT BY SIMILARITY.
FT GLYCERALDEHYDE 3-PHOSPHATE (BY
FT SIMILARITY).
FT ACTIVATES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
SQ SEQUENCE 335 AA; 35739 MW; PE7ACDFD7663E46 CRC64;

Query Match 66.2%; Score 1553.5; DB 1; Length 335;
Best Local Similarity 76.0%; Pred. No. 9.5e-88;
Matches 319; Conservative 2; Mismatches 14; Indels 85; Gaps 1;

QY 29 VVKVINGFGRIGRLAFRRIONIEGVETRIINDLTDPNMLAHLKYDITQGRFDGTVK 88
Db 1 VVKVINGFGRIGRLAFRRIONIEGVETRIINDLTDPNMLAHLKYDITQGRFDGTVK 60
QY 89 EGGFEVNGNFIKVSAREDPENIDWATDGVIEVLEALEGTVKDGDFVNGKFIKVSAREK 148
Db 61 EGGFEVNGNFIKVSAREDPENIDWATDGVIEVLEALEGTVKDGDFVNGKFIKVSAREK 80
QY 149 DPEQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLEA 208
Db 81 -----NIDWATDGVIEVLEA 95
QY 209 TSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVPNTNHDILDTGTVISGASCTTNC 268
Db 96 TGFPAKKEAEKHLHANGAKKVVITAPGNDVKTVPNTNHDILDTGTVISGASCTTNC 155
QY 269 LAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAANIVPNSGAAK 328
Db 156 LAPMAKALHDAFGIQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAANIVPNSGAAK 215
QY 329 AIGLVIPELNGKLDGAAQRPVPTGTVTLVTLTKVNSVDEINAAKAAASNDSEGYTED 388
Db 216 AIGLVIPELNGKLDGAAQRPVPTGTVTLVTLTKVNSVDEINAAKAAASNDSEGYTED 275
QY 389 PIVSSDIVGVSGLFDATQTKMVEDGSQLVKVSWYDNEMSYTAOLVRLTLEYFAKIAK 448
Db 276 PIVSSDIVGVSGLFDATQTKMVEDGSQLVKVSWYDNEMSYTAOLVRLTLEYFAKIAK 335

RESULT 4
G3P_LACLA
ID G3P_LACLA STANDARD; PRT; 337 AA.
AC P52987;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP OR LLO559.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LMO230;
RX MEDLINE=95291425; PubMed=7773380;
RA Cancilla M.R., Hillier A.J., Davidson B.E.;
RT "Lactococcus lactis glyceraldehyde-3-phosphate dehydrogenase gene;
RT gap: further evidence for strongly biased codon usage in glycolytic
RT pathway genes.";
RL Microbiology 141:1027-1036(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=L1403;
RX MEDLINE=21235186; PubMed=11337471;

```


RA Bolorin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weisenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium *Lactococcus*
RL *lactis* ssp. *lactis* IL1403.";
RL Genome Res. 11:731-753 (2001).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer (by similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L36907; AAC41453.1; -;
CC EMBL; AE006290; AAK04657.1; -;
CC PIR; G86694; G86694.
CC HSP; F17721; IHDG.
CC InterPro; IPR000173; GAP_dhhydrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh; C; 1.
CC PRINTS; PR00078; G3PDHGRGNASE.
CC TIGRfam; TIGR01534; GAPDH-I.1.
CC PROSITE; PS00071; GAPDH; FALSE_NEG.
CC Glycolysis; Oxidoreductase; NAD; Complete proteome.
CC FT BINDING 152 152
CC (BY SIMILARITY).
CC FT ACT_SITE 179 179
CC ACTIVATES THIOL GROUP DURING CATALYSIS
CC (BY SIMILARITY).
CC FT CONFLICT 143 143
CC T -> S (IN REF. 1).
CC SQ SEQUENCE 337 AA; 36057 MW; 17B8C6AAEFF589D CRC64;

Query Match 56.2%; Score 1280; DB 1; Length 337;
Best Local Similarity 61.4%; Pred. No. 4.2e-71;
Matches 259; Conservative 31; Mismatches 46; Indels 86; Gaps 2;

QY 28 MVKVGINGFGRGLAFRIQNVGEVETRLNDTPNLAHLKYDTQGRFDGTVEV 87
DB 1 MVKVGINGFGRGLAFRIQNVGEVETRLNDTPNLAHLKYDTQGRFDGTVEV 60

QY 88 KEGGFVNGNFVKYSAERDENIDWATDGVETVLEALEGTVEKGGFDVNGKFIKVSAA 147
DB 61 KEGGFVNGNFVKYSAERDENIDWATDGVETVLEALEGTVEKGGFDVNGKFIKVSAA 81

QY 148 KOPEQIDWATDGVETVLEIDGTVEVKEGGFVNGQFVKYSAERDENIDWATDGVETVLE 207
DB 82 -----DIQWADSGVEIVLE 95

QY 208 ATSPFAKAEAKHLHANGAKKVVITAPGNDVKTVVFNTHNLDLGTETVIGASCTTN 267
DB 96 ATGFATKAEAKHLHFGGAKKVVITAPGNDVKTVVFNTHNLDLGTETVIGASCTTN 155

QY 268 CLAPMAKALHDAFGIQKGLMTTHAYTGDQVILDPHRRGGLRRAGAAANIVPNSTGAA 327
DB 156 SLAPMADALNKNFVGKGTMTVHSYTGQDQVILDPHRRGGLRRAGAAANIVPNSTGAA 215

QY 328 KAIGVIVPELNGKLDGAQRVPTGVTETVLTLDKNVSVDEINAMKAASNDSEGYTE 387
DB 216 KAIGVIVPELNGKLDGAQRVPTGVTETVLTLDKNVSVDEINAMKAASNDSEGYTE 275

QY 388 DPIVSSDIIVGYSGLSDATQKMEV - DGSQLVKVVSVYDNEMSYTAQLVRLTEVPK 446
DB 276 DPIVSSDIIVGYSGLSDATQKMEV - DGSQLVKVVSVYDNEMSYTAQLVRLTEVPK 335

QY 447 AK 448

Db 336 AK 337

RESULT 5
GIP_CLOPA STANDARD; PRT; 334 AA.
ID Q59309;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH) (CP
DE 17/CP 18).
GN GAP.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]
SEQUENCE FROM N.A.
RA Oster T., Assobhei O., Scherrer S., Branlant G., Branlant C.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP PARTIAL SEQUENCE OF 1-26.
RC STRAIN=M5.
RX MEDLINE=98291870; PubMed=9629918;
RA Flengsrud R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5.";
RL Electrophoresis 19:802-806(1998).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X72219; CAAS1020.1; -;
CC PIR; S34254; S34254.
CC HSP; P00362; IGD1.
CC InterPro; IPR000173; GAP_dhhydrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh; C; 1.
CC PRINTS; PR00078; G3PDHGRGNASE.
CC TIGRfam; TIGR01534; GAPDH-I.1.
CC PROSITE; PS00071; GAPDH; 1.
CC Glycolysis; Oxidoreductase; NAD.
CC FT BINDING 150 150
CC ACTIVATES THIOL GROUP DURING CATALYSIS.
CC FT ACT_SITE 177 177
CC SEQUENCE 334 AA; 36078 MW; D15905D0DA7F62E7 CRC64;

Query Match 54.2%; Score 1234; DB 1; Length 334;
Best Local Similarity 59.3%; Pred. No. 2.6e-68;
Matches 249; Conservative 34; Mismatches 51; Indels 86; Gaps 3;

QY 29 VVKVGINGFGRGLAFRIQNVGEVETRLNDTPNLAHLKYDTQGRFDGTVEV 88
DB 1 VVKVGINGFGRGLAFRIQNVGEVETRLNDTPNLAHLKYDTQGRFDGTVEV 60

QY 89 EGGFVNGNFVKYSAERDENIDWATDGVETVLEALEGTVEKGGFDVNGKFIKVSAA 148
DB 61 EGGFVNGNFVKYSAERDENIDWATDGVETVLEALEGTVEKGGFDVNGKFIKVSAA 84


```
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE016745; AAC04154.1; -.
CC InterPro; IPR006424; GAPDH-I.
CC InterPro; IPR000173; GAP_dhdrogenase.
CC Pfam; PF00044; gpdh; 1.
CC PRINTS; PR00078; G3PDHGNASE.
CC TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
CC Glycolysis; Oxidoreductase; NAD; Complete proteome.
CC BINDING 151 151
CC ACT_SITE 178 178
CC (BY SIMILARITY).
CC (BY SIMILARITY).
CC SEQUENCE 336 AA; 36190 MW; A962202D02AB5767 CRC64;
CC
CC Query Match 50.0%; Score 1138.5; DS 1; Length 336;
CC Best Local Similarity 55.7%; Pred. No. 1.7e-62;
CC Matches 235; Conservative 39; Mismatches 61; Indels 87; Gaps 4;
CC
CC QY 28 MVKVGINGRGRIGRAFRIONVEGVETRIINDLPNMLHLLKYDTQREGDGTVEV 87
CC Db 1 MAIKVAINGRGRIGRAFRIONVEGVETRIINDLPNMLHLLKYDTQREGDGTVEV 80
CC
CC QY 88 KEGFVNGNFIKVSARDPENTDWDGVEIVLEALEGTVEVKDGGFDVNGKFIKVS 147
CC Db 61 IEGFRVNGKEIKSFDEPDAGKLPW-----GGLD----- 89
CC
CC QY 148 KDPEQIDWDGVEIVLEIDGTVEVKEGGEVNGQFVKVSAEPANIDWDGVEIVLE 207
CC Db 90 -----IDVLE 95
CC
CC QY 208 ATSPFAKKAERAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDTETVSGASCTTN 267
CC Db 96 CTGYTTDKRAQAHAIDA-GAKKVLISAPAKGVKTIVFNTNHDILDTGSETVSGASCTTN 154
CC
CC QY 268 CLAPMAKALHDAFGIOKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAIVNPNSTGAA 327
CC Db 155 SLAPVAKVLSDFGLVEGFMTHAYTGDQNTQDAPHRKGDGKRARAAAENIIPNSTGAA 214
CC
CC QY 328 KAIGLVPELNGKLDGAARVPVPTGVTVELVTLDK-NYSVDEINAAKKAASNDSPGYT 386
CC Db 215 KAIGKVIPEIDGKLDGAARVPVATGSLTSLTVLDKQDVTVQVNSAMKQASDESFGYT 274
CC
CC QY 387 EDPVTSVDIVGVSYSGLFDATQTKVMEVDGSQLVKVSVYDNEMSYTAQLVRLLEYFAKI 446
CC Db 275 EDEIVSSDIVGMTVGSLSFDATQTEVMTVGRQLVKVAAWYDNEMSYTAQLVRLTALHAE 334
CC
CC 447 AK 448
CC 335 SK 336
CC
CC RESULT 8
CC G3P1 STAA
CC ID G3P1 STAA
CC AC Q925C5;
CC DT 15-MAR-2004 (Rel. 43, Created)
CC DT 15-MAR-2004 (Rel. 43, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Glyceraldehyde 3-phosphate dehydrogenase 1 (EC 1.2.1.12) (GAPDH 1).
CC GN GAFA OR GAP OR SA0772 OR SA0727 OR MW0734.
```

```
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315),
OS Staphylococcus aureus (strain MW2), and
OS Staphylococcus aureus.
OX Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 158880, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BB;
RA Morrissey J.A., Williams P.;
RT "Isolation and characterization of a glycolytic operon in
RT Staphylococcus aureus";
RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-U. Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus";
RL Lancet 357:1225-1240(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwano M., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA";
RL Lancet 359:1819-1827(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwano M., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA";
RL Lancet 359:1819-1827(2002).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ133520; CAB38645.1; -.
CC EMBL; AF003360; BAB56934.1; -.
CC EMBL; AF003361; BAB41960.1; -.
CC EMBL; AF004824; BAB94599.1; -.
CC PIR; E89850; E89850.
CC SWISS-2DPAGE; Q925C5; STAA.
CC HSP; P17721; 1HDG.
CC InterPro; IPR006424; GAPDH-I.
CC InterPro; IPR000173; GAP_dhdrogenase.
CC Pfam; PF00044; gpdh; 1.
CC TIGRFAMs; TIGR01534; GAPDH-I; 1.
CC PROSITE; PS00071; GAPDH; 1.
CC Glycolysis; Oxidoreductase; NAD; Complete proteome.
CC BINDING 151 151
CC ACT_SITE 178 178
CC (BY SIMILARITY).
CC (BY SIMILARITY).
CC SEQUENCE 336 AA; 36281 MW; 37A6CEA9376779E5 CRC64;
CC
```


modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL: AJ000339; CAA04014.1; -
 DR PIR; T09633; T09633.
 DR HSP; P17721; LHDG.
 DR InterPro; IPR000173; GAP_dhdrogenase.
 DR InterPro; IPR006424; GAPDH-1.
 DR Pfam; PF00044; gpdh; 1.
 DR Pfam; PF02800; gpdh; 1.
 DR PRINTS; PR00078; G3PDHGRNASE.
 DR TIGRfam; TIGR01534; GAPDH-1.
 DR PROSITE; PS00071; GAPDH; FALSE NEG.
 DR GlycoSite; NAB; Oxidoreductase.
 KW BINDING 156 156 GLYCERALDEHYDE 3-PHOSPHATE
 (BY SIMILARITY).
 FT ACT_SITE 183 183 ACTIVATES THIOL GROUP DURING CATALYSIS
 (BY SIMILARITY).
 FT SEQUENCE 338 AA; 36564 MW; DF97E8C9CA4F7DFA CRC64;

Query Match 41.9%; Score 955.5; DB 1; Length 338;
 Best Local Similarity 47.3%; Pred. No. 2.3e-51;
 Matches 200; Conservative 46; Mismatches 88; Indels 89; Gaps 4;

QY 28 MVKVGINGFGRIQRLAFRRRI-----QNVGEGVEYTRINDLTDPNMLAHLKYDTTQGRFDG 83
 DB 1 MTWKIGINGFGRIQRLAFRRINDLGEETKDIEVVAINDLTTPAMLAHLKYDSTHGTFDH 60

QY 84 TVVKEGGFEVNGNFIVKSAERDPENIDWATDGVIEVLEALEGTVEVKGDFVNGKFK 143
 DB 61 EVSATEDSLVDGKKYRVVAEPOAQNPW-----VKN----- 92

QY 144 VSAEKPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGV 203
 DB 93 -----DQVD 96

QY 204 IVLEATSFPAKKEAEKHLHANGAKKVVITAPGNDVKTVPFNTHDILDGTETVIGAS 263
 DB 97 FVLECTGFYTSKAKSOAHLDA-GAKRVLISAPAGNDLKTIVSVNQDTLTADTTIVSAGS 155

QY 264 CTTNCLAPMAKALHDAFGQKGLMTTHAYTGQMTLDGPHRGDRLRRAGANIVPNS 323
 DB 156 CTTNSLAPMANALNKEFGIQVGTMTTHAYTATQKVLGDGDRGNFRNFAAENIIPHS 215

QY 324 TGAAKAIGLVIPELNGKLDGAAQRFVPTGVTGSELVTVVTLDKXNSVDPEINAAKASNDSP 383
 DB 216 TGAAKAIGLVLPELNGKLDGHAQRFVVDGSELTILDKKVTAEVNAAMKKVSPSF 275

QY 384 GYTEDPIVSDIVGVSGLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRLLEYF 443
 DB 276 AYNADQIVSTDVLGMTAGSIFDPTQTQITAGDKQLVKTVVAMVDYDNEYSFTQVMVRLHF 335

QY 444 AKI 446
 DB 336 ATL 338

RESULT 10
 G3P3_EC057
 ID_G3P3_EC057 STANDARD; PRT; 333 AA.
 AC P58072; 2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase C (EC 1.2.1.12) (GAPDH-C).
 GN GAPC OR Z2304 OR ECS2022.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 ON NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.

Query Match 49.7%; Score 1132.5; DB 1; Length 336;
 Best Local Similarity 54.7%; Pred. No. 3.9e-62;
 Matches 231; Conservative 44; Mismatches 60; Indels 87; Gaps 4;

QY 28 MVKVGINGFGRIQRLAFRRRIQNVGEGVEYTRINDLTDPNMLAHLKYDTTQGRFDGTV 87
 DB 1 MAVKVAINGFGRIQRLAFRRIOEVEGLEVVAVNDLTDMMLAHLKYDTTQGRFTGEVEV 60

QY 88 KEGGFEVNGNFIVKSAERDPENIDWATDGVIEVLEALEGTVEVKGDFVNGKFKVSAE 147
 DB 61 VDGGFRVNGKEVKSFESEFADSKLPWK-----DLN----- 89

QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEVLE 207
 DB 90 -----IDVLE 95

QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVPFNTHDILDGTETVIGASCTTN 267
 DB 96 CTGFYTDKDKQAHEA-GAKKVLISAPATGDLKTIIVTNHQLDGSSETVVGASCTTN 154

QY 268 CLAPMAKALHDAFGQKGLMTTHAYTGQMTLDGPHRGDRLRRAGANIVPNSGAA 327
 DB 155 SLAPVAKVLDNDFGLVEGLMTTHAYTGQNTQDAPKRGDKRRARAAENIIPNSGAA 214

QY 328 KAIGLVIPELNGKLDGAAQRFVPTGVTGSELVTVVTLDK-NVSVDEINAAKASNDSEY 386
 DB 215 KAIGKVIPEIDKLDGGAAQRFVPTGVTGSELVTVVTLDEKQDVTEVQVNEAMKVASNESFY 274

QY 387 EDPVSSDIVGVSGLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRLLEYFAKI 446
 DB 275 EDEIVSSDVGNTYGLSDATQTRVMSVGRQLVKVAAWYDNEMSYTAQLVRLLEYLAEL 334

QY 447 AK 448
 DB 335 SK 336

RESULT 9
 G3P3_LACDE
 ID_G3P3_LACDE STANDARD; PRT; 338 AA.
 AC O32755;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GAP.
 OS Lactobacillus delbrueckii (subsp. bulgaricus).
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 ON NCBI_TaxID=1585;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=107;
 RX MEDLINE=98240227; PubMed=9579064;
 RA Branny P., Delatorre F., Garel J.R.;
 RT "An operon encoding three glycolytic enzymes in Lactobacillus
 delbrueckii subsp. bulgaricus: glyceraldehyde-3-phosphate
 dehydrogenase, phosphoglycerate kinase and triosephosphate
 isomerase".
 RL Microbiology 144:905-914(1998).
 CC -|- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -|- PATHWAY: Second phase of glycolysis; first step.
 CC -|- SUBUNIT: Homotetramer (By similarity).
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -|- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 dehydrogenase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way


```

RC STRAIN=0157:H7 / BDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206651;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dmalanta E.F., Potamoussis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN=0157:H7 / RMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Tida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR ENBL; AE005364; ARG56359.1; -
DR ENBL; AP002557; EAB35445.1; -
DR PIR; C85737; C85737.
DR PIR; F90881; F90881.
DR HSP; P17721; IHDG.
DR InterPro; IPR00173; GAP dhhdrogenase.
DR InterPro; IPR006424; GAPDH-I.
DR Pfam; PF00044; Spdh; 1.
DR Pfam; PF02800; Spdh_C; 1.
DR PRINTS; PR00078; GPDHHRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
FT BINDING 150 150
FT ACT_SITE 177 177
FT ACTIVATES THIOL GROUP DURING CATALYSIS
FT (BY SIMILARITY).
FT ACT_SITE 333 AA; 35763 MW; A2F77CB2E773E64C CRC64;
FT (BY SIMILARITY).
Query Match 41.3%; Score 941.5; DB 1; Length 333;
Best Local Similarity 47.5%; Pred. No. 1.6e-50;
Matches 199; Conservative 53; Mismatches 76; Indels 91; Gaps 7;
QY 31 KVGNGFGRIGRLAPRIQNVQVEVTRINDLTPNNLAHLKYDTQRPFGTVEYKE 89
DB 3 KVGNGFGRIGRLVRLRLLEYSIDVVAINDLTPSKILAYLLKHSNYPGF----- 54
QY 90 GGFEVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVEYKGGFDVNGKFIKVSAREKD 149
DB 55 ----- 54
QY 150 PEQIDWATDGVIEVLEIDGTVEYKEGFEVNGQFKVSAREPANIDWATDGVIEVLEAT 209
DB 55 ---PWSVDFTEDSLIVD-----GKSIAYVAEKEAKNIPWKAAGAEIIVECT 97
QY 210 SFFAKKEAERKHLANGAKKVVITAPGNDGVKTVVFNNTNIDLTGTETVSGASCTTNCL 269

```

```

Db 98 GFYTSAEKSOAHLDA-GAKKVLISAPAG-EMKIVIVKVNDDTLDGNDTIIVSVASCTNCL 155
QY 270 APMAKALHDAFGIQGLMTTHIHYTQGMILDPGPHRGGLRRARAGAAANIVPNSGTGAAGA 329
Db 156 APMAKALHDSFGLEVGNTTHIHYTQSLVDGP-RGKDLRASRAAAENIIPHTTGAAGA 214
QY 330 IGLVPELNGKLDGAQRVPVPTGVTVELVTLTKVSVDEINAAKAA--SNDSEGYTE 387
Db 215 IGLVPELSGKLKGHAQRVPEVKTGSVTELVSILGKKVTAEEVNNALQKATNNESFGYTD 274
QY 388 DPVSSDIVGVSGLPFDATOTKVMEDVDSGLVKVWSYDNEMSYTAQLVRLTLEYFAKI 446
Db 275 EIVSSDLIISHFGSVFDAQTEITAVGDLQLVKTVWVYDNEVGFVTLRTLEKFAKL 333
RESULT 11
G3P3 ECOLI STANDARD; PRT; 333 AA.
AC P33898; P76094; P78062; P78291; Q03850; Q63208;
DT 01-FEB-1994 (Rel. 28, Created)
DT 26-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase C (EC 1.2.1.12) (GAPDH-C).
GN GAPC OR B1416/B1417.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=River isolate, and Clinical isolate;
RX MEDLINE=98283700; PubMed=9622357;
RA Espinosa-Urgel M., Kolter R.;
RT "Escherichia coli genes expressed preferentially in an aquatic
environment.";
RL Mol. Microbiol. 28:325-332(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Hidalgo E., Limon A., Aguilar J.;
RT "A second Escherichia coli gene with similarity to gapA.";
RL Microbiologia 12:99-106(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itch T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasubramanian S., Tagami H., Takeda J.,
RA Takeuchi K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-Kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [5]
RP SEQUENCE OF 306-332 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92011371; PubMed=1917845;
RA Hidalgo E., Chen Y.-M., Lin E.C.C., Aguilar J.;
RT "Molecular cloning and DNA sequencing of the Escherichia coli K-12
ald gene encoding aldehyde dehydrogenase.";

```


J. Bacteriol. 173:6118-6123 (1991).
 [6]
 RC SEQUENCE OF 1-13 FROM N.A.
 RC STRAIN-K12.
 RX MEDLINE=86232416; PubMed=2836696;
 RA Nakamura H., Murakami H., Yamato I., Anraku Y.;
 RT "Nucleotide sequence of the cydB gene encoding cytochrome b561 in
 Escherichia coli K12";
 RL Mol. Gen. Genet. 212:1-5 (1988).
 [7]
 RN SEQUENCE OF 1-165 FROM N.A.
 RA Krawetz S.A.;
 RL Submitted (OCT-1990) to the EMBL/GenBank/DBJ databases.
 [8]
 RN SEQUENCE OF 1-52 FROM N.A.
 RX MEDLINE=87053179; PubMed=3780374;
 RA Krawetz S.A., Connor W., Cannon P.D., Dixon G.H.;
 RT "A vector-primer-cloner-sequence plasmid for the construction of cDNA
 libraries: evidence for a rat glyceraldehyde-3-phosphate
 dehydrogenase-like mRNA and a ferritin mRNA within testis";
 RL DNA 5:427-435 (1986).
 CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate +
 NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -!- PATHWAY: Second phase of glycolysis; first step.
 CC -!- SUBUNIT: Homotetramer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 dehydrogenase family.
 CC -!- CAUTION: IN THE K12 STRAIN THIS GENE IS DISRUPTED BY A STOP CODON
 AND A FRAMESHIFT. IT SEEMS TO BE INTACT IN A NUMBER OF WILD
 STRAINS.
 CC -!- CAUTION: Ref.7 and Ref.8 sequences were originally thought to
 originate from rat.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L09067; AAA23956.1; ALT FRAME.
 DR EMBL; AB000239; AAC74498.1; ALT FRAME.
 DR EMBL; AB000239; AAC74499.1; ALT FRAME.
 DR EMBL; D90780; BAA15033.1; ALT FRAME.
 DR EMBL; D90781; BAA15038.1; ALT FRAME.
 DR EMBL; M64541; -; NOT ANNOTATED_CDS.
 DR EMBL; X07569; -; NOT ANNOTATED_CDS.
 DR EMBL; M4166; AAA41178.1; -.
 DR HSSP; P17721; 1HDG.
 DR EcoGene; EG12103; gapC.
 DR InterPro; IPR000173; GAP_dhhdhogenase.
 DR InterPro; IPR006424; GAPDH-I.
 DR Pfam; PF00044; gpdh; 1.
 DR Pfam; PF02800; gpdh; 1.
 DR TIGRfam; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 KW Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
 FT BINDING 150 150
 FT ACT_SITE 177 177
 FT ACTIVATES THIOL GROUP DURING CATALYSIS
 FT (BY SIMILARITY).
 FT (BY SIMILARITY).
 FT CONFLICT 39 39 K -> Y (IN REF. 7 AND 8).
 FT SEQUENCE 333 AA; 35649 MW; E88223297376B0A0 CRC64;
 Query Match 41.3%; Score 940.5; DB 1; Length 333;
 Best local similarity 47.5%; Pred. No. 1.9e-50;
 Matches 199; Conservative 53; Mismatches 76; Indels 91; Gaps 7;
 31 KVGINGFGRIGRILAFRIQVNE-GVEVTRINDLTPNMLAHLKYDTTQGRPDGTVEVKE 89

Db 3 KVGINGFGRIGRILVGRLLVKSNIIDVAINDLTSPKILAYILKHSNYGPF----- 54
 QY 90 GGFVANGNFIKVSARDPENIDWATDGEIVLEALEGTVEVKGDFVNGKFIKVSAREKD 149
 Db 55 ----- 54
 QY 150 PEQIDWATDGEIVLEIDGTVEVKEGFGFVNGOFVKVSAERSBPANIDWATDGDGVEIVLEAT 209
 Db 55 -----PMSVDFTEDSLVD-----GKSIVAYAEKAKNIPWAKAGAEIIVECT 97
 QY 210 SFFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDGTETVISGASCTTNCL 269
 Db 98 GFYTSAEKSAQALDA-GAKKVLISAPAG-EMKTIIVNANDDTLDGNDTIVSVASCTTNCL 155
 QY 270 APMKALHDAFGIQGLMTTHIAYTGDMLDGPHEGGDLRRARAGAAANIVPNSTGAAGA 329
 Db 156 APMKALHDSFGFVGTMTTHIAYTGTQSLVDGP-RGKDLRARRAAENLIIPHTGAAGA 214
 QY 330 IGLVPELNGKLDGAAQRPVPTGVTGTVTLVVTLDKQVSVDEINAAKAA--SNDSEGYTE 387
 Db 215 IGLVPELNGKLDGAAQRPVPTGVTGTVTLVVTLDKQVSVDEINAAKAA--SNDSEGYTE 274
 QY 388 DPVSSDIVSVGSLPFDATQTKMVEVDGSLQVKKVSVWDNEMSYTAQLVRIEYFAKI 446
 Db 275 EEIVSSDIIGSHFGSVFDTATQTEITAVGDLQVKTVAWYDNEYGFVTQLIRLTLEKFAKL 333
 RESULT 12
 G3P MYCPN
 ID -G3P MYCPN STANDARD; PRT; 337 AA.
 AC P75358;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GAPA OR GAP OR MPN430 OR MP411.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 pneumoniae";
 RL Nucleic Acids Res. 24:4420-4449 (1996).
 CC -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
 CC -!- PATHWAY: Second phase of glycolysis; first step.
 CC -!- SUBUNIT: Homotetramer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 dehydrogenase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB000040; AAB96059.1; -.
 CC PIR; S73737; S73737.
 DR HSSP; P17721; 1HDG.
 DR InterPro; IPR000173; GAP_dhhdhogenase.
 DR InterPro; IPR006424; GAPDH-I.
 DR Pfam; PF00044; gpdh; 1.
 DR Pfam; PF02800; gpdh; 1.
 DR PRINTS; PR00078; G3PDHGRGNASE.
 DR PROSITE; PS00071; GAPDH; 1.
 DR Glycolysis; Oxidoreductase; NAD; Multigene family; Complete proteome.
 FT BINDING 150 150
 FT ACT_SITE 177 177
 FT ACTIVATES THIOL GROUP DURING CATALYSIS
 FT (BY SIMILARITY).
 FT (BY SIMILARITY).
 FT CONFLICT 39 39 K -> Y (IN REF. 7 AND 8).
 FT SEQUENCE 333 AA; 35649 MW; E88223297376B0A0 CRC64;
 Query Match 41.3%; Score 940.5; DB 1; Length 333;
 Best local similarity 47.5%; Pred. No. 1.9e-50;
 Matches 199; Conservative 53; Mismatches 76; Indels 91; Gaps 7;
 31 KVGINGFGRIGRILAFRIQVNE-GVEVTRINDLTPNMLAHLKYDTTQGRPDGTVEVKE 89


```

J. Bacteriol. 175:7918-7930(1993)
CC -I- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -I- PATHWAY: Second phase of glycolysis; first step.
CC -I- SUBUNIT: Homotetramer (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL; U39710; RAC71523.1; --
CC DR EMBL; U02213; AAD12507.1; --
CC DR EMBL; U02178; AAD12463.1; --
CC DR PIR; C64233; C64233.
CC DR HSPG; P17721; IHDG.
CC DR TIGR; MG301; --
CC DR InterPro; IPR000173; GAP dhhydrogenase.
CC DR InterPro; IPR006424; GAPDH-I.
CC DR Pfam; PF00044; gpdh; 1.
CC DR Pfam; PF02800; gpdh; C; 1.
CC DR PRINTS; PR00078; G3PDHGRGNASE.
CC DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
CC DR PROSITE; PS00071; GAPDH; 1.
CC DR Glycolysis; Oxidoreductase; NAD; Complete proteome.
CC KW BINDING 157 157
CC FT FT FT
CC FT ACT_SITE 184 184
CC FT FT FT
CC FT SEQUENCE 337 AA; 37097 MW; FALEA196687006B CRC64;
CC -----
Query March 37.5%; Score 854.5; DB 1; Length 337;
Best Local Similarity 42.8%; Pred. No. 3.2e-45;
Matches 179; Conservative 52; Mismatches 99; Indels 89; Gaps 5

```

[illegible]


```

AC Q01651;
AD 01-APR-1993 (Rel. 25, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
DN GAP OR CGL1588.
GN Corynebacterium glutamicum (Brevibacterium flavum).
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
OC [1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 13059 / AS019;
RX MEDLINE=93015645; PubMed=1400158;
RA Eikmanns B.U.;
RT "Identification, sequence analysis, and expression of a
RT Corynebacterium glutamicum gene cluster encoding the three glycolytic
RT enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate
RT kinase, and triosephosphate isomerase.";
RL J. Bacteriol. 174:6076-6086(1992).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX Nakagawa S.;
RA "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X59403; CAA42045.1; -.
DR EMBL; AF005279; BAB98981.1; -.
DR DDB; A43260; A43260.
DR HSP; P00362; ICD1.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR InterPro; IPR006424; GAPDH-I.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; 1.
DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRFRAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
DR Glycolysis; Oxidoreductase; NAD; Complete proteome.
DR BINDING 153 153 GLYCERALDEHYDE 3-PHOSPHATE.
DR ACT SITE 180 180 ACTIVATES THICOL GROUP DURING CATALYSIS.
DR CONFLICT 25 26 SD -> NG (IN REF.1).
DR CONFLICT 333 334 KL -> QALN (IN REF.1).
DR SEQUENCE 334 AA; 36045 MW; 33792AF65FA90FA7 CRC64;
SQ
Query Match 37.3%; Score 850; DB 1; Length 334;
Best Local Similarity 44.7%; Pred. No. 6e-45;
Matches 189; Conservative 41; Mismatches 95; Indels 98; Gaps 7;

Qy 28 MVKVGINGFGRIGRLAFRI-QNVGVEVTRINDLTPNMLAHLKYDTQGRFDGTV 86
Db 1 MTRVINGFGRIGRNFRAVLERSDDLVAVVAVNDLTNKTLLTFLLXFDIMGRLGQVE 60
Qy 87 VKEGGFVNGNFIVKSAERDPENIDWATDGVIEIVLEAGTEVKDGGFDVNGKFIKVA 146
Db 61 YDDDSITVVGKRIAYAEIPDK----- 82
Qy 147 EKQPEQIDWATDGVIEIVLEIDGTVEVKEGGFVNGQFVKVSAERPANIDWATDGVIEIVL 206

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 29, 2004, 16:25:29 ; Search time 68.7761 Seconds
(without alignments)
2055.251 Million cell updates/sec

Title: US-10-650-369-22

Perfect score: 2278
Sequence: 1 MKKTGITILLAVILSAC.....EMSYAQLVRLTYFAKIAK 448

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1566.5	72.7	336	2 Q83ZF7	Q83zf7 streptococc
2	1564.5	68.7	336	2 Q9AJW2	Q9ajw2 streptococc
3	1564.5	68.7	336	16 Q8E3S8	Q8e3s8 streptococc
4	1564.5	68.7	336	16 Q8DXS8	Q8dxs8 streptococc
5	1559.5	68.5	336	2 Q8XVU6	Q8xvu6 streptococc
6	1544.5	67.8	336	2 Q8GCR7	Q8gcr7 streptococc
7	1535	67.4	335	16 Q97NL1	Q97nl1 streptococc
8	1535	67.4	359	16 Q8CWN6	Q8cwn6 streptococc
9	1534.5	67.4	336	2 Q8XHG1	Q8xhg1 streptococc
10	1530.5	67.2	336	2 Q8KVU5	Q8kvu5 streptococc
11	1521	66.8	337	16 Q8DVV3	Q8dvv3 streptococc
12	1513.5	66.4	336	2 Q8VVB9	Q8vvb9 streptococc
13	1459.5	64.1	320	2 Q9L5X6	Q9l5x6 streptococc
14	1410	61.9	309	2 Q9AJT7	Q9ajt7 streptococc
15	1409	61.9	309	2 Q9AJT9	Q9ajt9 streptococc
16	1407	61.8	309	2 Q9AJT4	Q9ajt4 streptococc

17	1405	61.7	309	2 Q9AJT5	Q9ajt5 streptococc
18	1404	61.6	308	2 Q9AJT8	Q9ajt8 streptococc
19	1387	59.9	305	2 Q9AJT6	Q9ajt6 streptococc
20	1357.5	59.6	336	16 Q9CDH4	Q9cdh4 lactococcus
21	1282	56.3	333	16 Q833I8	Q833i8 enterococcu
22	1205.5	52.9	336	16 Q8Y4I1	Q8y4i1 listeria mo
23	1204.5	52.9	336	16 Q828H9	Q828h9 listeria in
24	1175	51.6	332	16 Q8XKT9	Q8xkt9 clostridium
25	1164	51.1	335	16 Q8RFN9	Q8rfn9 fusobacteri
26	1145	50.3	334	16 Q9JTW8	Q9jtw8 neisseria m
27	1138.5	50.0	336	16 Q8CPY5	Q8cpy5 staphylococ
28	1134	49.8	330	2 Q83UX3	Q83ux3 neisseria g
29	1132.5	49.7	336	16 Q8Z5C5	Q8z5c5 staphylococ
30	1132	49.7	334	16 Q9UX95	Q9ux95 neisseria m
31	1131	49.6	330	2 Q84H26	Q84hz6 neisseria g
32	1130	49.6	330	2 Q84HZ4	Q84hz4 neisseria g
33	1129	49.6	335	2 Q93M61	Q93m61 staphylococ
34	1127	49.5	330	2 Q84HZ5	Q84hz5 neisseria g
35	1089	47.8	310	2 Q8L2P7	Q8l2p7 staphylococ
36	1082.5	47.5	311	2 Q8L2R0	Q8l2r0 staphylococ
37	1075	47.2	310	2 Q8L2Q2	Q8l2q2 staphylococ
38	1074	47.1	310	2 Q8L2Q6	Q8l2q6 staphylococ
39	1073	47.1	310	2 Q8L2P6	Q8l2p6 staphylococ
40	1071	47.0	310	2 Q8L2P5	Q8l2p5 staphylococ
41	1070	47.0	311	2 Q8VM79	Q8vm79 rhodospirill
42	1069	46.9	310	2 Q8L2Q3	Q8l2q3 staphylococ
43	1068.5	46.9	311	2 Q8L2P8	Q8l2p8 staphylococ
44	1067	46.8	310	2 Q8L2Q7	Q8l2q7 staphylococ
45	1065	46.8	310	2 Q8L2R7	Q8l2r7 staphylococ

ALIGNMENTS

RESULT 1

Q83ZF7 PRELIMINARY; PRT; 336 AA.
AC Q83ZF7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-P dehydrogenase.
GN GAPC.
OS Streptococcus dysgalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1334;
RN [1]
RP SEQUENCE FROM N.A.
RA Perez-Casal J.F., Bolton A., Song X.-M., Willson P., Potter A.A.;
RT "Use of the surface proteins GapC and Mig of Streptococcus
RT dysgalactiae as protective antigens against mastitis in non-lactating
RT cows".
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375662; AAF31408.1; -
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . . IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhhydrogenase.
DR Pfam; PF00044; spdh; 1.
DR Pfam; PF02800; spdh; 1.
DR PRINTS; PR00078; G3PDHGRNAS.
DR TIGRfams; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
SQ SEQUENCE 336 AA; 35928 MW; 11828218CF037076 CRC64;

Query Match 72.7%; Score 1656.5; DB 2; Length 336;
Best Local Similarity 79.8%; Pred. No. 2.3e-88;
Matches 335; Conservative 0; Mismatches 1; Indels 85; Gaps 1;

Qy 28 MVKVGINGFGRIGRLAFRIQNVGVETRIINDLTDPNMLAHLKYDTTQGRFDGTVVEV 87
Db 1 MVKVGINGFGRIGRLAFRIQNVGVETRIINDLTDPNMLAHLKYDTTQGRFDGTVVEV 60


```

QY 88 KEGGFVNGNFIKVAERDPENIDWATDGVIEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147
DB 61 KEGGFVNGNFIKVAERDPE-----81
QY 148 KDPQIDWATDGVIEIVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEIVLE 207
DB 82 -----NIDWATDGVIEIVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTWVNTNHDILDTGTTVISGASCTTN 267
DB 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTWVNTNHDILDTGTTVISGASCTTN 155
QY 268 CLAPMAKALHDADFQIOKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
DB 156 CLAPMAKALHDADFQIOKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKAAASNDSTGYTE 387
DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKAAASNDSTGYTE 275
QY 388 DPVSSDIVGVSGSLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
DB 276 DPVSSDIVGVSGSLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
DB 336 K 336

```

RESULT 2

```

Q9ALW2 PRELIMINARY; PRT; 336 AA.
ID Q9ALW2 AC Q9ALW2 DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J48;
RA Seifert K.N., Bleiweis A.S., McArthur W.P., Brady L.J.;
RT "The Group B Streptococcal Surface Antigen Delta is a Glyceraldehyde
RL 3-Phosphate Dehydrogenase."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF338416; AAK14387.1; -.
DR HSP; P00362; 1GDI.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhydrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

```

```

Query Match 68.7%; Score 1564.5; DB 2; Length 336;
Best Local Similarity 74.1%; Pred. No. 5e-83;
Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;
QY 28 MVKVGINGFGRIGRLAFRIQNVGEVETRIINDLPNMLHLKLYDTTQGRFDGTVEV 87
DB 1 MVKVGINGFGRIGRLAFRIQNVGEVETRIINDLPNMLHLKLYDTTQGRF-----54

```

```

QY 88 KEGGFVNGNFIKVAERDPENIDWATDGVIEIVLEALEGTVEVKDGGFDVNGKFIKVSAAE 147
DB 55 -----54
QY 148 KDPQIDWATDGVIEIVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEIVLE 207
DB 55 -----DGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEIVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTWVNTNHDILDTGTTVISGASCTTN 267
DB 96 ATGFFAKKEAAEKHLHANGAKKVVITAPGNDVKTWVNTNHDILDTGTTVISGASCTTN 155
QY 268 CLAPMAKALHDADFQIOKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 327
DB 156 CLAPMAKALQDNFVGKGLMTTHAYTGDMILDGPHRGDLRRARAGAAANIVPNSTGAA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKAAASNDSTGYTE 387
DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVSVELVTLTKNVSVDEINAAKAAASNDSTGYTE 275
QY 388 DPVSSDIVGVSGSLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
DB 276 DPVSSDIVGVSGSLFDATQTKWMEVDGSQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335
QY 448 K 448
DB 336 K 336

```

RESULT 3

```

Q9E3E8 PRELIMINARY; PRT; 336 AA.
ID Q9E3E8 AC Q9E3E8 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase.
GN G8S1811.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NM316 / Serotype III;
RX MEDLINE=2242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Prangeul L.,
RA Msadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease."
RL Mol. Microbiol. 45:1499-1513(2002).
DR EMBL; AL766853; CAD47470.1; -.
DR SAGAList; gbs1811; -.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhydrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; C; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Complete proteome.
SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

```

```

Query Match 68.7%; Score 1564.5; DB 16; Length 336;
Best Local Similarity 74.1%; Pred. No. 5e-83;
Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;
QY 28 MVKVGINGFGRIGRLAFRIQNVGEVETRIINDLPNMLHLKLYDTTQGRFDGTVEV 87
DB 1 MVKVGINGFGRIGRLAFRIQNVGEVETRIINDLPNMLHLKLYDTTQGRF-----54

```


Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRF----- 54
Qy 88 KEGGFVNGNFVKVSAERDPENIDWATDGVVEIVLEAGTVEVKGDFVNGKFIKVSAAE 147
Db 55 ----- 54
Qy 148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGFFVNGQVVKVSAERBPANIDWATDGVVEIVLE 207
Db 55 ----- 54
Qy 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTEVVISGASCTTN 267
Db 96 ATGFFASKEKAEQIHENGAKKVVITAPGNDVKTVFNTNHDILDTGTEVVISGASCTTN 155
Qy 268 CLAPMAKALHDAFGIOKGLMTTHAYTGDMILDPHGGDLRRARAGAAANIVPNSGTAA 327
Db 156 CLAPMAKALQDNFVKGQGLMTTHAYTGDMILDPHGGDLRRARAGAAANIVPNSGTAA 215
Qy 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLCKNVSVDEINAAKAAANDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVATLEKDVVEEVNAAKAAANDSFGYTE 275
Qy 388 DPVSSDIVGVSGLFDPATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVTLLEYFAKIA 447
Db 276 DPVSSDIVGVSGLFDPATQTKVQVDGNQLVKVSQWYDNEMSYTSQLVTLLEYFAKIA 335
Qy 448 K 448
Db 336 K 336

RESULT 4
Q8DXS8 PRELIMINARY; PRT; 336 AA.
ID Q8DXS8
AC Q8DXS8
DT 01-WAR-2003 (TREMELrel. 23, Created)
DT 01-WAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase.
GN GAP OR SAG1768.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RC MEDLINE=2222988; PubMed=12200547;
RA Tettelin H., Masignani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.W., Seaman M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carthy H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Malone D.,
RA Pinaudo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.,
RA "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
DR EMBL; AB014272; AAN0631.1; -;
DR TIGR; SAG1768; -;
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Complete proteome.
SQ SEQUENCE 336 AA; 36005 MW; 982E18A05CA343C9 CRC64;

Query Match 68.7%; Score 1564.5; DB 16; Length 336;
Best Local Similarity 74.1%; Pred. No. Se-83;
Matches 312; Conservative 15; Mismatches 9; Indels 85; Gaps 1;
Qy 28 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRFDTGV 87
Db 1 MVVKVINGFGRIGRLAFRRIONVEGVETRLNDLTPNMLAHLKYDTTQGRF----- 54
Qy 88 KEGGFVNGNFVKVSAERDPENIDWATDGVVEIVLEAGTVEVKGDFVNGKFIKVSAAE 147
Db 55 ----- 54
Qy 148 KDPEQIDWATDGVVEIVLEIDGTVEVKEGFFVNGQVVKVSAERBPANIDWATDGVVEIVLE 207
Db 55 ----- 54
Qy 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTEVVISGASCTTN 267
Db 96 ATGFFASKEKAEQIHENGAKKVVITAPGNDVKTVFNTNHDILDTGTEVVISGASCTTN 155
Qy 268 CLAPMAKALHDAFGIOKGLMTTHAYTGDMILDPHGGDLRRARAGAAANIVPNSGTAA 327
Db 156 CLAPMAKALQDNFVKGQGLMTTHAYTGDMILDPHGGDLRRARAGAAANIVPNSGTAA 215
Qy 328 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVTLCKNVSVDEINAAKAAANDSFGYTE 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGSGVTELVATLEKDVVEEVNAAKAAANDSFGYTE 275
Qy 388 DPVSSDIVGVSGLFDPATQTKVMEVDGSQLVKVSWYDNEMSYTAQLVTLLEYFAKIA 447
Db 276 DPVSSDIVGVSGLFDPATQTKVQVDGNQLVKVSQWYDNEMSYTSQLVTLLEYFAKIA 335
Qy 448 K 448
Db 336 K 336

RESULT 5
Q8KVU6 PRELIMINARY; PRT; 336 AA.
ID Q8KVU6
AC Q8KVU6
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
OS Streptococcus uberis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxID=1349;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 9927;
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,
RA Potter A.A.;
RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP
RT antigens confers protection against heterologous challenge with
RT Streptococcus uberis.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF421900; AAM73771.1; -;
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; 1.
DR PRINTS; PR00078; G3PDHDRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.

SQ SEQUENCE 336 AA; 35906 MW; 755D74A4548E61D4 CRC64;
 Query Match 68.5%; Score 1559.5; DB 2; Length 336;
 Best Local Similarity 73.9%; Pred. No. 9.7e-83;
 Matches 311; Conservative 15; Mismatches 10; Indels 85; Gaps 1;

QY 28 MVKVGINGGRIQLAFRIQNVGEVTRINDLTDNNLAHLKLYDTTQGRFPGTVEV 87
 DB 1 MVKVGINGGRIQLAFRIQNVGEVTRINDLTDNNLAHLKLYDTTQGRFPGTVEV 60

QY 88 KEGGFVNGNFIKVSAREDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSAE 147
 DB 61 KEGGFVNGNFIKVSAREDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSAE 81

QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 207
 DB 82 -----NIDWATDGVIEVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTTFVNTNHDILDTGTTVSGASCTTN 267
 DB 96 ATGFFAKKAAAEKHLHANGAKKVVITAPGDDVKTTFVNTNHDILDTGTTVSGASCTTN 155

QY 268 CLAPMAKALHDARIGIOKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAA 327
 DB 156 CLAPMAKALQDNFVGKGLMTTHAYTGDQMLDGPGRGDLRRARAGASNIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLVTLKXVSVDEINAAKAAANDSPGYTE 387
 DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLVTLKXVSVDEINAAKAAANDSPGYTE 275

QY 388 DPVSSDIVGVSGLFDATQTKMEVDGSLQVVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
 DB 276 DPVSSDIIGMAGSYLFDATQTKVQTVDGNQLVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448
 DB 336 K 336

RESULT 6
 Q8GCR7 PRELIMINARY; PRT; 336 AA.
 ID Q97NLI1
 AC Q97NLI1
 DT 01-OCT-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12).
 OS Streptococcus suis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1307;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S735;
 RA Brassard J., Gottschalk M., Quessey S.;
 RT "Cloning and purification of Streptococcus suis serotype 2
 glycerinaldehyde-3-phosphate dehydrogenase."
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY157026; AA896058.1;
 DR GO: GO:0004385; F:glyceraldehyde-3-phosphate dehydrogenase (p...); IEA.
 DR GO: GO:0001691; F:oxidoreductase activity; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR006424; GAPDH-I.
 DR InterPro: IPR000173; GAP_dhdrogenase.
 DR Pfam: PF00044; gpdh.1.
 DR Pfam: PF02800; gpdh.C1.
 DR PRINTS: PR00078; G3PDHGRGNASE.
 DR TIGRFAMs: TIGR01534; GAPDH-I; 1.
 DR PROSITE: PS00071; GAPDH; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 336 AA; 35925 MW; 1DBB9B1A492DCF59 CRC64;

Query Match 67.8%; Score 1544.5; DB 2; Length 336;

Best Local Similarity 72.9%; Pred. No. 7.2e-82;
 Matches 307; Conservative 16; Mismatches 13; Indels 85; Gaps 1;

QY 28 MVKVGINGGRIQLAFRIQNVGEVTRINDLTDNNLAHLKLYDTTQGRFPGTVEV 87
 DB 1 MVKVGINGGRIQLAFRIQNVGEVTRINDLTDNNLAHLKLYDTTQGRFPGTVEV 54

QY 88 KEGGFVNGNFIKVSAREDPENIDWATDGVIEVLEALEGTVEVKDGGFDVNGKFIKVSAE 147
 DB 55 ----- 54

QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 207
 DB 55 -----DGTVEVKEGGFVNGKFIKVSAREEPGNIDWATDGVIEVLE 95

QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKTTFVNTNHDILDTGTTVSGASCTTN 267
 DB 96 ATGFFAKKAAAEKHLHANGAKKVVITAPGNDVKTTFVNTNHDILDTGTTVSGASCTTN 155

QY 268 CLAPMAKALHDARIGIOKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAA 327
 DB 156 CLAPMAKALHDARIGIOKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLVTLKXVSVDEINAAKAAANDSPGYTE 387
 DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGSELVTLVTLKXVSVDEINAAKAAANDSPGYTE 275

QY 388 DPVSSDIVGVSGLFDATQTKMEVDGSLQVVKVSWYDNEMSYTAQLVRLTLEYFAKIA 447
 DB 276 DPVSSDIIGVIFSGSLFDATQTKVIEVDGSLQVVKVSWYDNEMSYTAQLVRLTLEYFAKIA 335

QY 448 K 448
 DB 336 K 336

RESULT 7
 Q97NLI1 PRELIMINARY; PRT; 335 AA.
 ID Q97NLI1
 AC Q97NLI1
 DT 01-OCT-2001 (TRENBLrel. 18, Created)
 DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12).
 GN SP2012 OR GAPDH.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC BAA-334 / TIGR4;
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Arguoli S., Dickinson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 pneumoniae";
 RL Science 293:498-506(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC11733;
 RA Bergmann S., Hammerschmidt S.;
 RT "Identification of pneumococcal GAPDH as plasmin(ogen)-binding
 protein";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF007490; AAK76079.1; -.
 DR EMBL: AJ505822; CAD44376.1; -.

PIR: F95235; F95235.
 DR HSSP: P00354; 3GPD.
 DR TIGR: SP2012; -.
 DR GO: GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p...; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR000424; GAPDH-I.
 DR InterPro: IPR000173; GAP_dhhydrogenase.
 DR Pfam: PF00044; gpdh; 1.
 DR Pfam: PF02800; gpdh; 1.
 DR PRINTS: PR00078; G3PDHGNASE.
 DR TIGRFAMs: TIGR01534; GAPDH-I.
 DR PROSITE: PS00071; GAPDH; 1.
 DR Oxidoreductase; Complete proteome.
 KW SEQUENCE 335 AA; 35856 MW; DR483CEA423E747B CRC64;
 SQ
 Query Match 67.4%; Score 1535; DB 16; Length 335;
 Best Local Similarity 73.6%; Pred. No. 2.5e-81;
 Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;
 QY 28 MVKVGINGFGRIGLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTV 87
 DB 1 MVKVGINGFGRIGLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTV 60
 QY 88 KEGGFVNGNFKVSAERDPENIDWATDGVIELEAGTVEVKDGGFDVNGKFIKVS 147
 DB 61 KEGGFVNGKFIKVSARDPE----- 81
 QY 148 KDPEQIDWATDGVIELEIDGTVEVEKGGFVNGQFVKVSAEREPANIDWATDGV 207
 DB 82 -----QIDWATDGVIEVLE 95
 QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTTVISGASCT 267
 DB 96 ATGFFAKKEAEKHL- KGGAKKVVITAPGNDVKTVFNTNHDVLDGTETVISGASCT 154
 QY 268 CLAPMAKALHDAFGTQKGLMTTHAYTGDMILDPHRRGGLRRARAGAAIVPNSTG 327
 DB 155 CLAPMAKALQDNFVGVVGLMTTHAYTGDMILDPHRRGGLRRARAGAAIVPNSTG 214
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGVTGVTGVTGVTGVTGVTGVTGVTGVT 387
 DB 215 KAIGLVIPELNGKLDGAAQRPVPTGVTGVTGVTGVTGVTGVTGVTGVTGVTGVT 274
 QY 388 DPIVSSDIVGSGSLFDATQTKWMEVDGSLQVKKVSWYDNEMSYTAQLVRLTFY 447
 DB 275 DPIVSSDIVGSGSLFDATQTKWMEVDGSLQVKKVSWYDNEMSYTAQLVRLTFY 334
 QY 448 K 448
 DB 335 K 335
 RESULT 8
 Q8CWN6 PRELIMINARY; PRT; 359 AA.
 ID Q8CWN6
 AC Q8CWN6;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (Phosphorylating)
 DE (SC 1.2.1.12).
 GN GAPD OR SPRI825.
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=171101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21429245; PubMed=11544234;
 RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
 RA DeHoff B.S., Bateman S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,

RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
 RA Sun P.-M., Winkler M.B., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
 RA Glass J.I.;
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6";
 RL J. Bacteriol. 183:5709-5717(2001).
 DR ENBL: AS008547; AAL00628.1; -.
 DR PIR: G98099; G98099.
 DR GO: GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p...; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR000173; GAPDH-I.
 DR InterPro: IPR000173; GAP_dhhydrogenase.
 DR Pfam: PF00044; gpdh; 1.
 DR Pfam: PF02800; gpdh; 1.
 DR PRINTS: PR00078; G3PDHGNASE.
 DR TIGRFAMs: TIGR01534; GAPDH-I.
 DR PROSITE: PS00071; GAPDH; 1.
 DR Oxidoreductase; Complete proteome.
 KW SEQUENCE 359 AA; 38763 MW; 61EF0E375B330B60 CRC64;
 SQ
 Query Match 67.4%; Score 1535; DB 16; Length 359;
 Best Local Similarity 73.6%; Pred. No. 2.8e-81;
 Matches 310; Conservative 12; Mismatches 13; Indels 86; Gaps 2;
 QY 28 MVKVGINGFGRIGLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTV 87
 DB 25 MVKVGINGFGRIGLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRFDGTV 84
 QY 88 KEGGFVNGNFKVSAERDPENIDWATDGVIELEAGTVEVKDGGFDVNGKFIKVS 147
 DB 85 KEGGFVNGKFIKVSARDPE----- 105
 QY 148 KDPEQIDWATDGVIELEIDGTVEVEKGGFVNGQFVKVSAEREPANIDWATDGV 207
 DB 106 -----QIDWATDGVIEVLE 119
 QY 208 ATSPFAKKEAEKHLHANGAKKVVITAPGNDVKTVFNTNHDILDTGTTVISGASCT 267
 DB 120 ATGFFAKKEAEKHL- KGGAKKVVITAPGNDVKTVFNTNHDVLDGTETVISGASCT 178
 QY 268 CLAPMAKALHDAFGTQKGLMTTHAYTGDMILDPHRRGGLRRARAGAAIVPNSTG 327
 DB 179 CLAPMAKALQDNFVGVVGLMTTHAYTGDMILDPHRRGGLRRARAGAAIVPNSTG 238
 QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGVTGVTGVTGVTGVTGVTGVTGVTGVT 387
 DB 239 KAIGLVIPELNGKLDGAAQRPVPTGVTGVTGVTGVTGVTGVTGVTGVTGVTGVT 298
 QY 388 DPIVSSDIVGSGSLFDATQTKWMEVDGSLQVKKVSWYDNEMSYTAQLVRLTFY 447
 DB 299 DPIVSSDIVGSGSLFDATQTKWMEVDGSLQVKKVSWYDNEMSYTAQLVRLTFY 358
 QY 448 K 448
 DB 359 K 359
 RESULT 9
 Q8KHG1 PRELIMINARY; PRT; 336 AA.
 ID Q8KHG1
 AC Q8KHG1;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Glyceraldehyde-3-phosphate dehydrogenase.
 OS Streptococcus agalactiae, and
 OS Streptococcus infant.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1311, 1346;
 RN [1]

RP SEQUENCE FROM N.A.
RC SPECIES=S.agalactiae, and S.iniae; STRAIN=ATCC 27541, and 9117;
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,
RA Potter A.A.;
RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP
RT antigens confers protection against heterologous challenge with
RT Streptococcus uberis";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF421899; AAM73770.1; -;
DR EMBL; AF421902; AAM73773.1; -;
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; C; 1.
DR PRINTS; PRO0078; G3PDHGRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 35723 MW; ABAB9EL4F3ED1111 CRC64;

Query Match 67.4%; Score 1534.5; DB 2; Length 336;
Best Local Similarity 72.7%; Pred. No. 2.7e-81;
Matches 306; Conservative 17; Mismatches 13; Indels 85; Gaps 1;

QY 28 MVKVGINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRPDGVEV 87
DB 1 MVKVGINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRPDGVEV 60

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALGTVVEKDGDFVNGKFIKVS 147
DB 61 KDGGFVNGS-----GTVEVKGDFVNGKFIKVS 70

QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEV 207
DB 71 -----FVKVSAEREPANIDWATDGVIEV 95

QY 208 ATSPFAKKAABKHANGAKKVVITAPGNDVKTVVTNTNHDILDTGTTVISGASCTTN 267
DB 96 ATGFPAAKAAEOKHANGAKKVVITAPGNDVKTVVTNTNHDILDTGTTVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIOKGLMTTIHAYTGDMILDGPHRGDLRRARAGANNIVPNSGAA 327
DB 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDMILDGPHRGDLRRARAGANNIVPNSGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLVTLTKNVSVDENNAKKAASNDSPGYTE 387
DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLVTLVAVLEKTSVEENNAKKAASNDSPGYTE 275

QY 388 DPVSSDIVGVSGSLFDATQTKMVEVDGSQLKVVSWYDNEMSYTAQLVRLTYFAKIA 447
DB 276 DAIVSSDIVGVSGSLFDATQTKVTVGDNQVLKVVSWYDNEMSYTAQLVRLTYFAKIA 335

448 K 448
336 K 336

RESULT 10
Q8KVU5 PRELIMINARY; PRT; 336 AA.
ID Q8KVU5
AC Q8KVU5
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
OS Streptococcus parvauberis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

OX NCBI_TaxID=1348;
RN [1]
RP SEQUENCE FROM N.A.
RA Fontaine M.C., Perez-Casal J., Song X.-M., Shelford J., Willson P.J.,
RA Potter A.A.;
RT "Immunization of dairy cattle with recombinant GapC and chimeric CAMP
RT antigens confers protection against heterologous challenge with
RT Streptococcus uberis";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF421901; AAM73772.1; -;
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR GO; GO:0004365; F:oxidoreductase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; C; 1.
DR PRINTS; PRO0078; G3PDHGRGNASE.
DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 36037 MW; 311647C25489AC9E CRC64;

Query Match 67.2%; Score 1530.5; DB 2; Length 336;
Best Local Similarity 73.4%; Pred. No. 4.6e-81;
Matches 309; Conservative 13; Mismatches 14; Indels 85; Gaps 2;

QY 28 MVKVGINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRPDGVEV 87
DB 1 MVKVGINGFGRIQLAFRRIONVEGVETRIINDLTPNMLAHLKYDTTQGRPDGVEV 55

QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALGTVVEKDGDFVNGKFIKVS 147
DB 56 -----GTVEVKGDFVNGKFIKVS 77

QY 148 KDPEQIDWATDGVIEVLEIDGTVEVKEGFEVNGQFVKVSAEREPANIDWATDGVIEV 207
DB 78 KDPEQ-----IDWATDGVIEV 95

QY 208 ATSPFAKKAABKHANGAKKVVITAPGNDVKTVVTNTNHDILDTGTTVISGASCTTN 267
DB 96 ATGFPAAKAAEOKHANGAKKVVITAPGNDVKTVVTNTNHDILDTGTTVISGASCTTN 155

QY 268 CLAPMAKALHDAFGIOKGLMTTIHAYTGDMILDGPHRGDLRRARAGANNIVPNSGAA 327
DB 156 CLAPMAKALQDNFVGKQGLMTTIHAYTGDMILDGPHRGDLRRARAGANNIVPNSGAA 215

QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLVTLTKNVSVDENNAKKAASNDSPGYTE 387
DB 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLVTLVAVLEKTSVEENNAKKAASNDSPGYTE 275

QY 388 DPVSSDIVGVSGSLFDATQTKMVEVDGSQLKVVSWYDNEMSYTAQLVRLTYFAKIA 447
DB 276 DPVSSDIVGVSGSLFDATQTKVTVGDNQVLKVVSWYDNEMSYTAQLVRLTYFAKIA 335

448 K 448
336 K 336

RESULT 11
Q8DVV3 PRELIMINARY; PRT; 337 AA.
ID Q8DVV3
AC Q8DVV3
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Extracellular glyceraldehyde-3-phosphate dehydrogenase (EC
DE 1.2.1.12).
OS Streptococcus mutans.
OS Streptococcus mutans.

OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1308;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LW1811;
RA van den Bogaard P.T.C., Kleerebezem M., Hols P., Crispie F.,
RA Kuipers O.P., de Vos W.M.,
RT "Modulation of Glycolysis by Lactose Availability in Streptococcus
RT thermophilus,"
RL Submitted (Nov-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF442551; AAJ35377.1; --
DR HSPF; P00354; 36PD.
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; C; 1.
DR PRINTS; PR00078; G3FDHGRNASE.
DR TIGRfams; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
DR Oxidoreductase.
KW Oxidoreductase.
SQ SEQUENCE 336 AA; 36026 MW; 52C1F25F3A7E0230 CRC64;
Query Match 66.4%; Score 1513.5; DB 2; Length 336;
Best Local Similarity 71.7%; Pred. No. 4.5e-80;
Matches 302; Conservative 18; Mismatches 16; Indels 85; Gaps 1;
QY 28 MVKVGINGGRIGRLAFRRIONVEGVETRLNDLPNNLAHLKLYDTTQGRFDGTVVEV 87
Db 1 MVKVGINGGRIGRLAFRRIONVEGVETRLNDLPNNLAHLKLYDTTQGRFDGTVVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVVEKDGDFVNGKFIKVSAR 147
Db 61 KEGGFVNGNFIKVSARDPE----- 81
QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 82 -----QIDWANDGVIEVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKVTWFTNTHDILDTGTETVISGASCTTN 267
Db 96 ATGFTTKVLAEKHLHPGAKKVVITAPGNDVKVTWFTNTHDILDTGTETVISGASCTTN 155
QY 269 CLAPMAKALHDAPGIGKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGA 327
Db 156 CLAPMAKALHDNFIKEGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGA 215
QY 328 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLVDKNVSVDEINAAKKAASNDSPFGYT 387
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLVDKNVSVDEINAAKKAASNDSPFGYT 275
QY 388 EDPVSSDIVGVSYSGLFDATQTKVMEVDGSQLVKVSVYDNEMSYTAQLVRLTLEYPAKI 447
Db 276 EDPVSSDIVGVSYSGLFDATQTKVLDVGDGKLVKVVSYDNEMSYTSQLVRLTLEYPAKI 335
QY 448 AK 448
Db 336 AK 336
RESULT 13
Q9LSX6 PRELIMINARY; PRT; 320 AA.
ID Q9LSX6
AC Q9LSX6
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RA MEDLINE=22295063; PubMed=12397186;
RA Aidic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.,
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen,"
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
DR EMBL; A5014883; AAN58118.1; --
DR GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. .; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; C; 1.
DR PRINTS; PR00078; G3FDHGRNASE.
DR TIGRfams; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
DR Oxidoreductase; Complete proteome.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 337 AA; 36068 MW; 42BF20365963C22 CRC64;
Query Match 66.8%; Score 1521; DB 16; Length 337;
Best Local Similarity 71.8%; Pred. No. 1.7e-80;
Matches 303; Conservative 18; Mismatches 15; Indels 86; Gaps 2;
QY 28 MVKVGINGGRIGRLAFRRIONVEGVETRLNDLPNNLAHLKLYDTTQGRFDGTVVEV 87
Db 1 MVKVGINGGRIGRLAFRRIONVEGVETRLNDLPNNLAHLKLYDTTQGRFDGTVVEV 60
QY 88 KEGGFVNGNFIKVSARDPENIDWATDGVIEVLEALEGTVVEKDGDFVNGKFIKVSAR 147
Db 61 KEGGFVNGNFIKVSARDPE----- 81
QY 148 KDPQIDWATDGVIEVLEIDGTVEVKEGGFVNGQFVKVSAEREPANIDWATDGVIEVLE 207
Db 82 -----QIDWANDGVIEVLE 95
QY 208 ATSPFAKKEAAEKHLHANGAKKVVITAPGNDVKVTWFTNTHDILDTGTETVISGASCTTN 266
Db 96 ATGFTTKVLAEKHLHPGAKKVVITAPGNDVKVTWFTNTHDILDTGTETVISGASCTTN 155
QY 267 NCLAPMAKALHDAPGIGKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGA 326
Db 156 NCLAPMAKALHDNFIKEGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIVPNSTGA 215
QY 327 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLVDKNVSVDEINAAKKAASNDSPFGYT 386
Db 216 KAIGLVIPELNGKLDGAAQRPVPTGVTGLVTLVDKNVSVDEINAAKKAASNDSPFGYT 275
QY 387 EDPVSSDIVGVSYSGLFDATQTKVMEVDGSQLVKVSVYDNEMSYTAQLVRLTLEYPAKI 446
Db 276 EDPVSSDIVGVSYSGLFDATQTKVLDVGDGKLVKVVSYDNEMSYTSQLVRLTLEYPAKI 335
QY 447 AK 448
Db 336 AK 337
RESULT 12
Q8VVB9 PRELIMINARY; PRT; 336 AA.
ID Q8VVB9
AC Q8VVB9
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde-3-phosphate dehydrogenase.
GN GAPDH.

DE Extracellular glycerolaldehyde-3-phosphate dehydrogenase
DE (Fragment).
GN Streptococcus gordonii.
OS Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RA Nelson D., Goldstein J.M., Boatright K., Hartly D.W.S., Cook S.L.,
RA Hickman P.J., Potempa J., Travis J., Mayo J.A.;
RT "Purification and characterization of an extracellular
RT glycerolaldehyde-3-phosphate dehydrogenase from Streptococcus sanguis
RT and cloning of the gene encoding this enzyme."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AF247678; AAF64063.1; -.
DR HSP; P00362; IGD1.
DR GO; GO:0004365; F:glycerolaldehyde-3-phosphate dehydrogenase (p...); IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; C; 1.
DR PRINTS; PR00078; G3PDHGRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
FT NON_TER 1
FT NON_TER 320
FT NON_TER 320
SQ SEQUENCE 320 AA; 33968 MW; 6FFCC18BB01E91F0 CRC64;
Query Match 64.1%; Score 1459.5; DB 2; Length 320;
Best Local Similarity 71.9%; Pred. No. 5.7e-77;
Matches 291; Conservative 14; Mismatches 15; Indels 85; Gaps 1;
QY 31 KVGINGFGRIGLAFRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEVKEG 90
DB 1 KVGINGFGRIGLAFRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEVKEG 60
QY 91 GFVNGNFIKVSARDENIDWATDGVLEIVLEALEGIVKDGFDVNGKFIKVSARKDP 150
DB 61 GFVNGKFKVSAERDPE-----78
QY 151 EQIDWATDGVLEIVLEIDGTVEVKEGFEVNGQFVKVSAERANIDWATDGVLEIVLEATS 210
DB 79 -----NIDWANDGVLEIVLEATG 95
QY 211 FPAKKEAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDTETVISGASCTTNCLPA 270
DB 96 FPAKAAAEKHLHAGGAKKVVITAPGGSDVKTIVFNTNHDVLDGTETVISGASCTTNCLPA 155
QY 271 PMAKALHDAFGQKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIPNSGAAKAI 330
DB 156 PMAKALQNFVGVGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIPNSGAAKAI 215
QY 331 GLVIFELNGKLDGAAQRPVPTGVTSLVTLWLDKNVSVDEINAAKAAASNDSPGYTDEPDI 390
DB 216 GLVIFELNGKLDGAAQRPVPTGVTSLVTLVAVLDKNVTVDEVNAAKAAASNESYGYTDEPDI 275
QY 391 VSSDIVGVYSGSLFATQTKMEVDGSQLVKVSWYDNEMSYTAQ 435
DB 276 VSSDIVGVYSGSLFATQTKVLDVVGKQLVKVSWYDNEMSYTAQ 320
RESULT 14
Q9AJT7 PRELIMINARY; PRT; 309 AA.
AC Q9AJT7
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Glycerolaldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
DE (Fragment).
GN GAP.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RA Amzaga M.R., Carter P.E., Cash P., McKenzie H.;
RA "Sequence variation in mef(A) and gap genes in M-phenotype
RT erythromycin resistant Streptococcus pneumoniae."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC DEHYDROGENASE FAMILY.
DR EMBL; AJ292048; CAC27448.1; -.
DR HSP; P00362; IGD1.
DR GO; GO:0004365; F:glycerolaldehyde-3-phosphate dehydrogenase (p...); IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR006424; GAPDH-I.
DR InterPro; IPR000173; GAP_dhdrogenase.
DR Pfam; PF00044; gpdh; 1.
DR Pfam; PF02800; gpdh; C; 1.
DR PRINTS; PR00078; G3PDHGRGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Oxidoreductase.
FT NON_TER 1
FT NON_TER 309
FT NON_TER 309
SQ SEQUENCE 309 AA; 32930 MW; 89E45C3BEA6A0528 CRC64;
Query Match 61.9%; Score 1410; DB 2; Length 309;
Best Local Similarity 71.9%; Pred. No. 4e-74; 13; Indels 86; Gaps 2;
Matches 284; Conservative 12; Mismatches 12; Indels 86; Gaps 2;
QY 36 GFGRIGRIGLAFRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEVKEGFEVN 95
DB 1 GFGRIGRIGLAFRIQNVGEVETRIINDLTPNMLAHLKYDTTQGRFDGTVEVKEGFEVN 60
QY 96 GNFIKVSARDENIDWATDGVLEIVLEALEGIVKDGFDVNGKFIKVSARKDPQIDW 155
DB 61 GNFIKVSARDPE-----73
QY 156 ATDGVLEIVLEIDGTVEVKEGFEVNGQFVKVSAERANIDWATDGVLEIVLEATGFFAKK 215
DB 74 -----QIDWATDGVLEIVLEATGFFAKK 95
QY 216 EAAEKHLHANGAKKVVITAPGGNDVKTIVFNTNHDILDTETVISGASCTTNCLAPMAKA 275
DB 96 EAAEKHL-KGGAKKVVITAPGGNDVKTIVFNTNHDVLDGTETVISGASCTTNCLAPMAKA 154
QY 276 LHDAPFOIKGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIPNSGAAKAIGLVIP 335
DB 155 LQDNFVGVGLMTTHAYTGDQMLDGPGRGDLRRARAGAAANIPNSGAAKAIGLVIP 214
QY 336 ELNGKLDGAAQRPVPTGVTSLVTLWLDKNVSVDEINAAKAAASNDSPGYTDEPDISSDI 395
DB 215 ELNGKLDGAAQRPVPTGVTSLVTLVAVLEKNTVDEVNAAKAAASNESYGYTDEPDISSDI 274
QY 396 VGVYSGSLFATQTKMEVDGSQLVKVSWYDNEM 430
DB 275 VGVYSGSLFATQTKVLDVVGKQLVKVSWYDNEM 309
RESULT 15
Q9AJT9 PRELIMINARY; PRT; 309 AA.
AC Q9AJT9
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

Wed Mar 31 10:09:17 2004

Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)
 DE (Fragment).
 DE GN GAP.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=clinical isolate;
 RA Amezaga M.R., Carter P.E., Cash P., McKenzie H.;
 RT "Sequence variation in mef(A) and gap genes in M-phenotype
 RT erythromycin resistant Streptococcus pneumoniae";
 EL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 DR EMBL; AJ232046; CAC27446.1; --
 DR HSP; P00362; IGL1.
 DR GO; GO:0004365; F:Glyceraldehyde-3-phosphate dehydrogenase (p. . .); IEA.
 DR GO; GO:0016491; F:Oxidoreductase activity; IEA.
 DR GO; GO:0006096; P:Glycolysis; IEA.
 DR InterPro; IPR006424; GAPDH-I.
 DR InterPro; IPR000173; GAP_dhhydrogenase.
 DR Pfam; PF00044; Spdh_1.
 DR Pfam; PF02800; Spdh_C; 1.
 DR PRINTS; PRO0078; G3PDHDRGNASE.
 DR TIGRFAMS; TIGR01534; GAPDH-I; 1.
 DR PROSITE; PS00071; GAPDH; 1.
 KW Oxidoreductase.
 FT NON_TER 1
 FT NON_TER 309
 SQ SEQUENCE 309 AA; 32916 MW; 39E9CC3BEA6F8959 CRC64;
 Query Match 61.9%; Score 1409; DB 2; Length 309;
 Best Local Similarity 71.6%; Pred. No. 4.6e-74;
 Matches 283; Conservative 13; Mismatches 13; Indels 86; Gaps 2;
 QY 36 GFGRIQLAFRRIONVEGVETRIINDLTPNNLAHLKYDTTQGRPDGTVEVKEGFEVN 95
 DB 1 GFGRIQLAFRRIONVEGVETRIINDLTPVLAHLKYDTTQGRPDGTVEVKEGFEVN 60
 QY 96 GNFIKVSAERDPENIDWATDGVLEALEGTVKVGDFVNGKFIKVSAREKDEQIDW 155
 DB 61 GKFKVSAERDPE-----QIDWATDGVLEALEGTVKVGDFVNGKFIKVSAREKDEQIDW 73
 QY 156 ATDGVLEIDGTVEVKEGFEVNGQFVKVSAERBPANIDWATDGVLEALEGTVKVGDFVNGKFIKVSAREKDEQIDW 215
 DB 74 -----QIDWATDGVLEALEGTVKVGDFVNGKFIKVSAREKDEQIDW 95
 QY 216 EAAEKHLHANGAKKVITAPGNDVKTVPNTNHDLDGTETVIGASCTTNCIAPMAKA 275
 DB 96 EAAEKHL-KGAKKVITAPGNDVKTVPNTNHDLDGTETVIGASCTTNCIAPMAKA 154
 QY 276 LHDAGTQKGLMTTHAYTGDMILDPHGGDLRRRAGAANTVPNSTGAAGAIGLVIP 335
 DB 155 LQDNFVGVEGLMTTHAYTGDMILDPHGGDLRRRAGAANTVPNSTGAAGAIGLVIP 214
 QY 336 ELNCKLDGAORVPVPTGVTVELVTLDKNVSDVEINAAKASNDISFGYTEDPIVSSDI 395
 DB 215 ELNCKLDGAORVPVPTGVTVELVTLDKNVSDVEINAAKASNDISFGYTEDPIVSSDI 274
 QY 396 VGSYGSILFDATQTKVMEVDGSQLVKVSYVDNEM 430
 DB 275 VGSYGSILFDATQTKVMEVDGSQLVKVSYVDNEM 309

Search completed: March 29, 2004, 16:35:17
 Job time : 70.7761 secs